

UNIVERSIDAD AUTÓNOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR



PROYECTO FIN DE CARRERA

**Implementación y comparación de soluciones basadas en Path
Computation Element (PCE) para entornos multi-dominio**

Diego Álvarez Martínez

Septiembre 2011

Implementación y comparación de soluciones basadas en Path Computation Element (PCE) para entornos multi-dominio

AUTOR: Diego Álvarez Martínez

TUTOR: Víctor López Álvarez

PONENTE: Jorge E. López de Vergara Méndez

High Performance Computing and Networking Research Group

Dpto. de Tecnología Electrónica y de las Comunicaciones

Escuela Politécnica Superior

Universidad Autónoma de Madrid

Septiembre de 2011

Resumen

La arquitectura Automatic Switched Optical Network (ASON) y Generalized Multi-Protocol Label Switching (GMPLS) son el resultado de grandes esfuerzos en los últimos años en la definición de plano de control para redes ópticas. Este modelo permite a los nodos de red realizar las tareas de cálculo de caminos de una manera distribuida. Sin embargo, cuando los algoritmos de cómputo son complejos, como ocurre en las redes ópticas, el cálculo de caminos se vuelve una tarea muy costosa. Esto aumenta los requisitos computacionales en los nodos de red y, por tanto, su coste.

La estandarización de la arquitectura Path Computation Element (PCE) surge con el objetivo de liberar a los nodos de red de las tareas de cálculo de caminos. En entornos multi-dominio existe un problema sobre cómo realizar el intercambio de información para obtener un cálculo óptimo de rutas, sin intercambiar información confidencial de cada dominio. La arquitectura PCE se presenta como una solución apropiada en el diseño de las redes multi-dominio actuales. Dicho PCE es la entidad encargada de calcular los caminos solicitados por los Path Computation Clients (PCC) utilizando la información almacenada en su Base de Datos de Ingeniería de Tráfico (Traffic Engineering Database - TED). Al ser una entidad centralizada, sólo dispone de información referente a su dominio, por lo que podrá colaborar con PCEs de otros dominios para calcular caminos multi-dominio óptimos. En la comunicación entre PCC y PCE, así como entre PCEs, utilizan el protocolo Path Computation Element Communication Protocol (PCEP).

En este proyecto se estudian distintas soluciones basadas en PCE para entornos multi-dominio: *Per-domain Path Computation*, *Backward-Recursive PCE-based Computation* y *PCE Jerárquico*. Se realiza una implementación del PCE para estas soluciones, se valida su funcionamiento y se analiza su comportamiento en términos de probabilidad de bloqueo mediante una serie de experimentos. Se proponen e implementan soluciones para la mejora de los resultados obtenidos y se determina su impacto mediante los mismos experimentos. Por último, se realiza una comparación entre las distintas soluciones y se concluyen los resultados, obteniendo así nuevas líneas de investigación futuras.

Palabras Clave

PCE, cálculo de caminos, multi-dominio, per-domain, BRPC, PCE Jerárquico, probabilidad de bloqueo

Abstract

The Automatic Switched Optical Network (ASON) architecture and Generalized Multi-Protocol Label Switching (GMPLS) are the result of great efforts in recent years in the definition of control plane for optical networks. This model allows network nodes to perform path computation tasks in a distributed way. However, when computation algorithms are complex, as it occurs in optical networks, path computation becomes a very expensive task. This increases the computational requirements in network nodes, and therefore their costs.

The standardization of the Path Computation Element (PCE) architecture was created with the aim of freeing network nodes of path computation tasks. In multi-domain network scenarios there is an issue about how to exchange information in order to obtain an optimum path computation, without exchanging confidential information of each domain. The PCE architecture is a suitable solution for the design of actual multi-domain networks. Such PCE is in charge of the computation of the paths requested by the Path Computation Clients (PCC) using information stored in its Traffic Engineering Database (TED). Since it is a centralized entity, it only has information about its domain, so it may cooperate with PCEs of other domains to calculate optimum multi-domain paths. In the communication between PCC and PCE, and between PCEs as well, they use the Path Computation Element Communication Protocol (PCEP).

This project studies different PCE-based solutions for multi-domain scenarios: *Per-domain Path Computation*, *Backward-Recursive PCE-based Computation* and *Hierarchical PCE*. It carries out a PCE implementation for these solutions, evaluates their performance and analyzes their behaviour in terms of blocking probability through a set of experiments. It proposes and implements solutions in order to improve the results and determines their impact through the same experiments. Finally, it carries out a comparison of the implemented solutions and concludes with the results, thus obtaining new future research lines.

Key Words

PCE, path computation, multi-domain, per-domain, BRPC, Hierarchical PCE, blocking probability

Agradecimientos

En primer lugar me gustaría agradecer enormemente a mi familia el esfuerzo que han hecho durante muchos años para darme una educación y unos valores que me han permitido afrontar todos los retos que se han presentado en esta etapa de mi vida. Gracias por apoyarme y disfrutar a mi lado de los buenos momentos, pero también por entender y soportar los no tan buenos. Sin vosotros todo esto no hubiera sido posible.

También me gustaría dar las gracias a mis compañeros de carrera, pero por encima de todo, amigos, Alex, Álvaro y Jose. Estos años han sido muy duros, pero al final lo hemos conseguido con esfuerzo y constancia. Ha sido un verdadero placer.

Gracias también a Rubén por estar siempre que le he necesitado, en los buenos y malos momentos, desde que teníamos cuatro años. Gracias, amigo.

A José Luis Añamuro, por su ayuda en los inicios de este proyecto.

Y por último, a Víctor López, mi tutor. Gracias por todo tu trabajo y esfuerzo durante este tiempo. Siempre has estado disponible cuando lo he necesitado, me has ayudado a conseguir los objetivos marcados y, sobre todo, has sabido transmitirme tu apoyo y confianza en todo momento.

A todos, gracias.

Índice de Contenidos

Resumen.....	I
Abstract.....	III
Agradecimientos	V
Índice de Contenidos	VII
Índice de Figuras.....	IX
Índice de Tablas.....	XIII
Glosario de Acrónimos.....	XV
1. Introducción.....	1
1.1. Motivación	2
1.2. Objetivos	3
1.3. Estructura.....	3
2. PCE en entornos multi-dominio.....	5
2.1. Path Computation Element (PCE)	6
2.1.1. Arquitectura del modelo PCE.....	7
2.1.2. Path Computation Element Communication Protocol (PCEP).....	12
2.2. Entornos Multi-dominio.....	17
2.2.1. Enrutamiento en redes multi-dominio	18
2.2.2. Señalización en redes multi-dominio	20
2.2.3. Problema del cálculo de caminos multi-dominio	20
2.2.4. Tipos de bloqueo para PCE en entornos multi-dominio	21
3. Soluciones basadas en PCE para entornos multi-dominio	23
3.1. Per-Domain Path Computation.....	23
3.2. Simple Cooperating PCEs	24
3.3. Backward-Recursive PCE-Based Computation (BRPC).....	26
3.4. PCE Jerárquico (H-PCE).....	29
4. Comparativa de las soluciones basadas en PCE para entornos multi-dominio	33
4.1. Descripción del escenario	33

4.2.	Probabilidad de bloqueo	40
4.3.	Soluciones propuestas para la mejora de la probabilidad de bloqueo por “robo de longitud de onda”	45
5.	Conclusiones y futuras líneas de trabajo	51
5.1.	Conclusiones.....	51
5.2.	Futuras líneas de trabajo	53
	Referencias.....	55
Anexo A.	Validación de resultados	59
Anexo B.	Manual de usuario	67
Anexo C.	Presupuesto.....	73
Anexo D.	Pliego de condiciones	75
Anexo E.	Publicación.....	81

Índice de Figuras

Figura 1.1 - Arquitectura básica del modelo PCE.....	1
Figura 1.2 - Colaboración entre PCEs en un escenario multi-dominio.....	2
Figura 2.1 - Red de Transporte de Próxima Generación.....	5
Figura 2.2 - PCE en el Plano de Control.....	7
Figura 2.3 - PCE en el Plano de Gestión	8
Figura 2.4 - Configuraciones conjunta (izda.) y separada (dcha.) del PCE	8
Figura 2.5 - Descripción funcional de la arquitectura basada en PCE.....	9
Figura 2.6 - Modelo <i>centralizado</i> de la arquitectura PCE.....	10
Figura 2.7 - Modelo <i>distribuido</i> de la arquitectura PCE	11
Figura 2.8 - Sesiones PCEP.....	13
Figura 2.9 – Notificaciones enviadas por el PCE (izda.) y el PCC (dcha.).....	14
Figura 2.10 – Mensaje de <i>Error</i> enviado por el PCE	15
Figura 2.11 – Escenario con múltiples dominios o ASes de diferentes operadores.....	17
Figura 2.12 - Espacios de direcciones IP separados	18
Figura 2.13 - Espacio de direcciones IP compartido.....	19
Figura 2.14 - Señalización en una red multi-dominio.....	20
Figura 2.15 - Escenario multi-dominio.....	20
Figura 3.1 – Per-domain Path Computation	23
Figura 3.2 – Simple Cooperating PCEs.....	25
Figura 3.3 – Problema de optimización en Simple Cooperating PCEs.....	25
Figura 3.4 – Backward-Recursive PCE-based Computation.....	26
Figura 3.5 – Escenario complejo para Backward-Recursive PCE-based Computation.....	28
Figura 3.6 – Arquitectura PCE Jerárquica.....	29
Figura 3.7 – PCE Jerárquico	31
Figura 4.1 - <i>Red Troncal China</i> [31]	33
Figura 4.2 - Composición de un PCE.....	34

Figura 4.3 – Generación de tráfico en un PCC Fuente.....	36
Figura 4.4 - Implementación del Per-domain Path Computation.....	37
Figura 4.5 - Implementación del BRPC	38
Figura 4.6 - Implementación del H-PCE	39
Figura 4.7 - Estabilidad para per-domain.....	40
Figura 4.8 - Estabilidad para BRPC	40
Figura 4.9 - Estabilidad para H-PCE.....	41
Figura 4.10 - Probabilidad de bloqueo para BRPC variando K con el algoritmo FF	41
Figura 4.11 - Tipos de bloqueo (izda.) y tiempos de cómputo (dcha.) para BRPC variando K con el algoritmo FF.....	42
Figura 4.12 - Probabilidad de bloqueo para H-PCE variando K con el algoritmo FF.....	43
Figura 4.13 - Tipos de bloqueo (izda.) y tiempos de cómputo (dcha.) para H-PCE variando K con el algoritmo FF.....	43
Figura 4.14 - Probabilidad de bloqueo para per-domain con el algoritmo FF.....	44
Figura 4.15 - Tipos de bloqueo para BRPC con los algoritmos FF (izda.) y RR-FF (dcha.).....	45
Figura 4.16 - Tipos de bloqueo para H-PCE con los algoritmos FF (izda.) y RR-FF (dcha.)	46
Figura 4.17 - Probabilidad de bloqueo para BRPC con los algoritmos FF y RR-FF y PR.....	47
Figura 4.18 - Tipos de bloqueo para BRPC con los algoritmos FF (izda.) y RR-FF (centro) y PR (dcha.).	47
Figura 4.19 - Probabilidad de bloqueo para H-PCE con los algoritmos FF y RR-FF y PR.....	48
Figura 4.20 - Tipos de bloqueo para H-PCE con los algoritmos FF (izda.), RR-FF (centro) y PR (dcha.).	48
Figura 4.21 - Comparativa en términos de probabilidad de bloqueo con el mecanismo PR.....	49
Figura 4.22 - Comparativa de los tiempos de cómputo con el mecanismo PR.....	50
Figura A.1 - Topología experimental <i>Ochos</i>	59
Figura A.2 - Estabilidad para per-domain.....	60
Figura A.3 - Estabilidad para BRPC	60
Figura A.4 - Estabilidad para H-PCE.....	61
Figura A.5 - Probabilidad de bloqueo para BRPC variando K con el algoritmo FF	61
Figura A.6 - Tipos de bloqueo para BRPC variando K con el algoritmo FF.....	62

Figura A.7 - Probabilidad de bloqueo para H-PCE variando K con el algoritmo FF	62
Figura A.8 - Tipos de bloqueo para H-PCE variando K con el algoritmo FF	63
Figura A.9 - Probabilidad de bloqueo para per-domain con el algoritmo FF	63
Figura A.10 - Probabilidad de bloqueo para BRPC con los algoritmos FF y RR-FF y PR.....	64
Figura A.11 - Tipos de bloqueo para BRPC con los algoritmos FF y RR-FF y PR.....	64
Figura A.12 - Probabilidad de bloqueo para H-PCE con los algoritmos FF y RR-FF y PR.....	65
Figura A.13 - Tipos de bloqueo para H-PCE con los algoritmos FF y RR-FF y PR.....	65
Figura A.14 - Comparativa en términos de probabilidad de bloqueo con el mecanismo PR	66
Figura B.1 - <i>Red Troncal China</i> [31].....	68
Figura B.2 - Ejecución del PCE 1 antes (arriba) y después (abajo)	69
Figura B.3 - Ejecución del PCC Fuente 1 con tasas deterministas (arriba) y exponenciales (abajo).....	70

Índice de Tablas

Tabla 2.1 - Tipos de Notificaciones del protocolo PCEP	15
Tabla 2.2 – Tipos de Errores del protocolo PCEP	16
Tabla 4.1 – Características de la <i>Red Troncal China</i> (36 nodos)	34
Tabla A.1 – Características de la topología experimental <i>Ochos</i> (24 nodos)	59

Glosario de Acrónimos

AS	Autonomous System
ASON	Automatically Switched Optical Network
BGP	Border Gateway Protocol
BRPC	Backward-Recursive PCE-based Computation
E-NNI	External Network to Network Interface
ERO	Exclude Route Object
FF	First-Fit
FTTX	Fiber To The X
GMPLS	Generalized Multi-Protocol Label Switching
H-PCE	Hierarchical Path Computation Element
HSPA	High Speed Packet Access
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
I-NNI	Internal Network to Network Interface
IP	Internet Protocol
IRO	Include Route Object
IS-IS	Intermediate System-Intermediate System
LER	Label Edge Router
LSA	Link State Advertisement
LSP	Label Switched Path
LSR	Label Switch Router
LTE	Long Term Evolution
MPLS	Multi-Protocol Label Switching
NGTN	Next Generation Transport Network
NMS	Network Management System
NNI	Network to Network Interface
OSPF	Open Shortest Path First
PCC	Path Computation Client
PCE	Path Computation Element

PCEP	Path Computation Element Communication Protocol
PFC	Proyecto Fin de Carrera
PR	Pre-Reserva
QoS	Quality of Service
RP	Request Parameters
RR-FF	Round-Robin First-Fit
RSVP	Resource reSerVation Protocol
RTT	Round-Trip Time
TCP	Transmission Control Protocol
TE	Traffic Engineering
TED	Traffic Engineering Database
TLV	Type-Length-Value
UMTS	Universal Mobile Telecommunications System
UNI	User to Network Interface
VSPT	Virtual Shortest Path First
WDM	Wavelength-Division Multiplexing
WiMAX	Worldwide Interoperability for Microwave Access

1. Introducción

En los últimos años, las redes de telecomunicaciones han experimentado un crecimiento exponencial en el tráfico de datos debido, en gran medida, a la expansión de Internet. La aparición de nuevos servicios demandantes de gran ancho de banda tales como la televisión en alta definición y en 3D, vídeo bajo demanda, juegos en red, cloud computing, videoconferencia, etc., unido a la implementación de nuevas tecnologías de acceso tanto fijas (redes FTTX activas y pasivas) como móviles (UMTS, HSPA, LTE o WiMAX) hace pensar que el tráfico seguirá aumentando en un futuro próximo.

Actualmente, los operadores de red están migrando a Redes de Transporte de Próxima Generación (Next Generation Transport Networks - NGTN) con el fin de ser capaces de proveer servicios de voz, datos y vídeo garantizando Calidad de Servicio (Quality of Service - QoS) [1]. Estas redes están basadas en el concepto “Topo IP” (All-IP), que trata de integrar todo tipo de servicios en una misma infraestructura de red IP, abaratando así los costes. Las redes troncales están formadas por enrutadores IP/MPLS y utilizan tecnología WDM [2].

Tradicionalmente las redes troncales se gestionaban de forma centralizada, de manera que los elementos de red eran configurados estáticamente. En estos últimos años se han hecho grandes esfuerzos en la creación de un plano de control común mediante la estandarización de Generalized Multi-Protocol Label Switching (GMPLS) [4], que proporciona configuración dinámica y distribuida de la capa óptica. Sin embargo, el cálculo de caminos en redes ópticas es una tarea compleja en términos de computación al tener en cuenta las restricciones adicionales de los elementos de red ópticos. Cuando dicha tarea deben realizarla los controladores GMPLS, se hace necesario contar con equipos con suficientes recursos computacionales y, por tanto, el coste de los mismos aumenta.

Es por ello que surge un nuevo elemento de las redes de nueva generación denominado Path Computation Element (PCE) [8] con el objetivo de liberar a los nodos de la red de las tareas de cómputo de caminos aplicando restricciones de Ingeniería de Tráfico (Traffic Engineering - TE). Dicha ingeniería de tráfico puede ser entendida como la capacidad para monitorizar y gestionar el comportamiento de las redes con el fin de sacar provecho de ellas y proporcionar calidad de servicio [3]. El PCE se define como una entidad de red que contiene información sobre la topología y es consultado por los nodos de la red para determinar el camino a seguir desde un origen hasta un destino. Así mismo, utiliza información abstracta del dominio al que pertenece para calcular caminos óptimos. Por tanto, el uso del PCE elimina la necesidad de computar rutas dentro de los nodos de la red (Figura 1.1), reduciendo su coste.

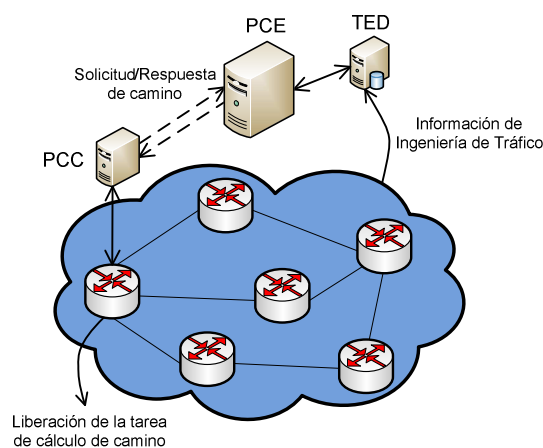


Figura 1.1 - Arquitectura básica del modelo PCE

En una arquitectura basada en el modelo PCE hay, al menos, un PCE en cada dominio. Sin embargo, un dominio puede contener múltiples PCEs para facilitar el balanceo de carga y evitar puntos únicos de fallo. El PCE recibe solicitudes de cálculo de caminos desde los Path Computation Clients (PCC). Para atenderlas, necesita información actualizada del estado de la red, la cual almacena en la Base de Datos de Ingeniería de Tráfico (Traffic Engineering Database - TED). El escenario de la Figura 1.1 muestra un ejemplo básico de este modelo.

Dicha arquitectura proporciona la funcionalidad necesaria para el cálculo de caminos óptimos con ingeniería de tráfico en redes GMPLS. Además de liberar a los nodos de la red de las tareas de cómputo, se debe asegurar que el tráfico llegue a su destino con seguridad, fiabilidad y optimizando los recursos de la red sobre la que viaja.

Al tratarse de una entidad centralizada, un PCE es capaz de calcular caminos óptimos de manera local (dentro de su dominio), pero también puede colaborar con otros PCEs de diferentes dominios en la consecución de caminos inter-dominio óptimos.

1.1. Motivación

Un PCE es capaz de computar LSPs (Label Switched Paths) sobre su propio dominio. Sin embargo, cuando el destino de la ruta solicitada no forma parte de su dominio, se hace necesaria la colaboración entre PCEs de diferentes dominios (véase la Figura 1.2). Dichos PCEs pueden cooperar de igual a igual o de manera jerárquica [9]. Cada uno es el responsable de calcular el segmento del camino perteneciente a su dominio, ya que dispone de la información necesaria para ello [9].

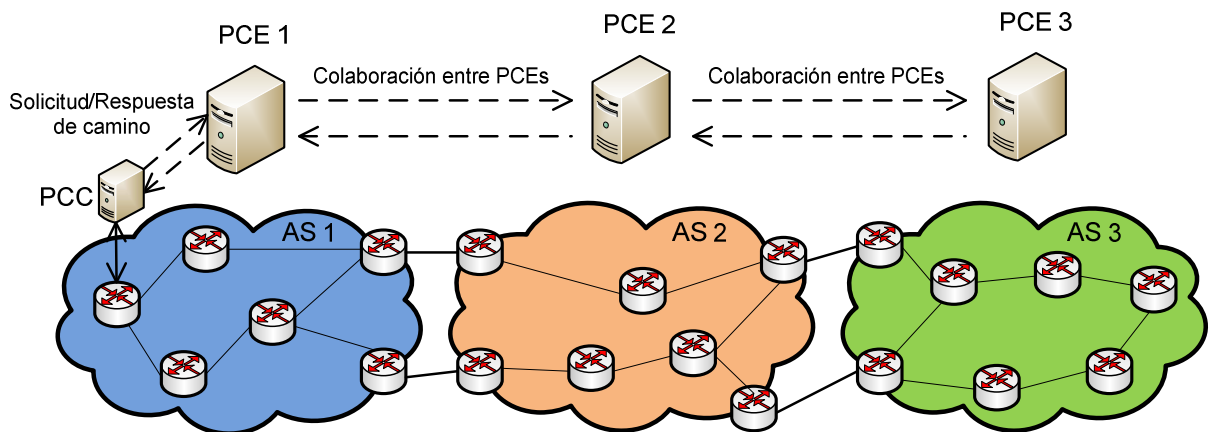


Figura 1.2 - Colaboración entre PCEs en un escenario multi-dominio

En este escenario, la arquitectura PCE debe contar con mecanismos para abordar los problemas que presenta el encaminamiento inter-dominio, tales como la visibilidad limitada de la topología de la red, la falta de información de ingeniería de tráfico o la confidencialidad de la información interna de los dominios.

Se han definido distintas soluciones para abordar este problema: *Simple Cooperating PCEs* [9], *Per-domain Path Computation* [21], *Backward-Recursive PCE-based Computation (BRPC)* [23] y *PCE Jerárquico (Hierarchical PCE - H-PCE)* [23]. Todas ellas se basan en la cooperación de los diferentes PCEs a lo largo de los dominios comprendidos entre el origen y el destino. Sin embargo, no se ha estudiado completamente su rendimiento en estos entornos multi-dominio.

1.2. Objetivos

La finalidad de este Proyecto Fin de Carrera (PFC) es dotar a una implementación del PCE de soporte para las distintas soluciones propuestas en entornos multi-dominio y comprobar su funcionamiento. En este proyecto se trabajará para conseguir los siguientes objetivos:

- Estudiar y analizar el estado del arte del PCE en entornos multi-dominio para redes ópticas.
- Implementar las extensiones necesarias para soportar las soluciones elegidas.
- Definir escenarios para comprobar el funcionamiento adecuado del PCE con soporte para las soluciones implementadas.
- Validar el funcionamiento y comparar las soluciones implementadas mediante una serie de experimentos.

1.3. Estructura

Esta memoria se divide en cinco capítulos:

- El primer capítulo es una introducción al problema estudiado. Se exponen la motivación y los objetivos de este proyecto.
- En el segundo capítulo se estudia en detalle la arquitectura basada en PCE y se explica el funcionamiento del protocolo de comunicaciones Path Computation Element Communication Protocol (PCEP). Además, se exponen las características principales de los entornos multi-dominio, tales como el enrutamiento y la señalización, y se presenta el problema del cálculo de caminos en dichos entornos multi-dominio, así como los tipos de bloqueo que se pueden dar.
- En el tercer capítulo se exponen las soluciones existentes para el cálculo de caminos en entornos multi-dominio: *Per-domain Path Computation*, *Simple Cooperating PCEs*, *Backward-Recursive PCE-based Computation* y *PCE Jerárquico*. Se explica su funcionamiento y se muestran las extensiones necesarias para su implementación.
- El cuarto capítulo define el escenario utilizado para la realización de los experimentos. Se presentan resultados en términos de probabilidad de bloqueo para las soluciones implementadas y se proponen alternativas a los problemas presentados. Por último, se analizan los resultados obtenidos y se realiza una comparación en términos de probabilidad de bloqueo para las soluciones implementadas.

- El quinto capítulo concluye este proyecto mediante la evaluación de los resultados obtenidos y propone nuevas líneas de trabajo futuras.
- Por último, se pueden encontrar una serie de anexos para completar el trabajo realizado. En el anexo A se muestran los resultados para un escenario multi-dominio distinto con el objetivo de contrastar los mismos. El anexo B es un manual de usuario, orientado a poder utilizar las implementaciones realizadas de una manera adecuada. El anexo C es el presupuesto del proyecto, el anexo D el pliego de condiciones y en el anexo E se presenta una publicación fruto del trabajo realizado en este PFC.

2. PCE en entornos multi-dominio

En este capítulo se presenta el modelo de cálculo de caminos basado en PCE y su aplicación a entornos multi-dominio. El modelo de red utilizado para su definición, así como de los planos de control y gestión se muestra en la Figura 2.1.

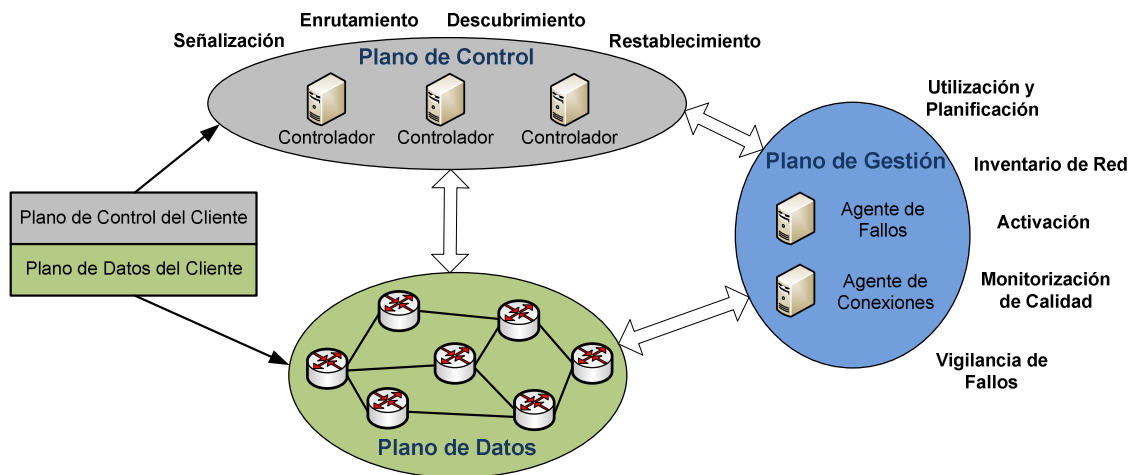


Figura 2.1 - Red de Transporte de Próxima Generación

El plano de datos se usa para la transmisión de los paquetes de información y está compuesto por enlaces IP/MPLS u ópticos por donde viajan los datos de los usuarios. El plano de gestión se ocupa de las operaciones globales, incluyendo la monitorización o la gestión de fallos. El plano de control se encarga de cuestiones de gestión descentralizada, tales como el intercambio de información de enrutamiento, el control del estado de los enlaces y el establecimiento y liberación de conexiones.

La aparición de estas redes de transporte de próxima generación plantea una incógnita: cómo establecer caminos entre múltiples dominios. La utilización conjunta de la arquitectura ASON [4][1] y GMPLS [4], junto con el modelo basado en PCE, permiten encontrar una solución a este problema.

Este capítulo se divide en dos secciones. En la sección 2.1 se explica en detalle el PCE, su arquitectura y su funcionamiento. En la sección 2.2 se presentan las soluciones existentes para el enrutamiento y la señalización en redes multi-dominio y se plantea el problema del cálculo de caminos revisando los tipos de bloqueo que se pueden presentar en una arquitectura basada en PCE para redes ópticas.

2.1. Path Computation Element (PCE)

El cálculo de caminos con restricciones es una característica fundamental para la ingeniería de tráfico en las redes MPLS [6] y GMPLS [7]. Este proceso resulta muy costoso cuando se trata de grandes redes multi-dominio y requiere la utilización de componentes computacionales especiales, además de la colaboración entre elementos de diferentes dominios.

Según el IETF, un PCE es *“una entidad (componente, aplicación o nodo de red) capaz de calcular una ruta basada en un grafo, aplicando restricciones computacionales”* [8]. Por tanto, la capacidad de cómputo de caminos en la red se sitúa en estos elementos especiales denominados PCEs.

Existen diferentes situaciones en las que es apropiado usar una arquitectura basada en PCE. Cabe destacar que este modelo no pretende sustituir a otros modelos existentes de cálculo de rutas, sino que es adecuado para determinados escenarios o situaciones que pueden presentarse en las redes actuales.

En [8] se detallan estas situaciones. Las siguientes son algunas de ellas:

- **Alto consumo de recursos:** puede ocurrir que el cálculo de una ruta resulte tan complejo que el Label Switch Router (LSR) no pueda llevarlo a cabo porque no dispone de suficientes recursos para ello. En este caso, dicha tarea pasará a ser responsabilidad de un PCE.
- **Visibilidad limitada:** existen multitud de situaciones en donde el LSR responsable del cálculo del camino no dispone de la información necesaria para llegar al destino. Es el caso de un escenario multi-dominio, en donde el LSR no tiene acceso a la información de otros dominios. Este problema puede ser solucionado mediante un modelo basado en PCE en donde los PCEs de diferentes dominios colaboran entre sí en el cálculo de la ruta.
- **Ausencia de TED:** mantener la TED puede requerir el uso de mucha memoria y una cantidad de recursos muy alta. Es posible también que algunos nodos dentro de un dominio no soporten las extensiones de ingeniería de tráfico de los protocolos de enrutamiento, por lo que se hace necesario que la TED sea suministrada por un PCE.
- **Ausencia de plano de control o capacidad de enrutamiento:** En el ámbito de las redes ópticas es común que un elemento de la red no posea ni plano de control ni capacidad de enrutamiento, por lo que todas sus conexiones son establecidas desde el plano de gestión. En este caso el PCC formaría parte del plano de gestión y sería el Sistema de Gestión de Red (Network Management System - NMS) el que solicitaría el cálculo de las rutas al PCE.
- **Cálculo de rutas alternativas:** se puede usar un PCE para calcular caminos alternativos cuando un Label Switched Path (LSP) necesite ser restablecido rápidamente o por motivos de seguridad.
- **Aplicación:** debe entenderse que la arquitectura basada en PCE no es una solución aplicable a grandes redes (por ejemplo, Internet), sino que su aplicación se encuentra en escenarios con una serie de dominios cuyas relaciones son conocidas.

2.1.1. Arquitectura del modelo PCE

La arquitectura básica de un modelo basado en PCE está compuesta por tres elementos fundamentales: Path Computation Client (PCC), Path Computation Element (PCE) y Base de Datos de Ingeniería de Tráfico (TED). En el ámbito de las redes MPLS y GMPLS, un PCC es un LSR (router MPLS/GMPLS) de la red que solicita el cálculo de un TE LSP a un PCE, y éste se encarga de calcularlo y de proporcionárselo al PCC. Para el cálculo de caminos, un PCE utiliza la información de ingeniería de tráfico contenida en la TED. En este caso el PCE forma parte del plano de control, véase la Figura 2.2.

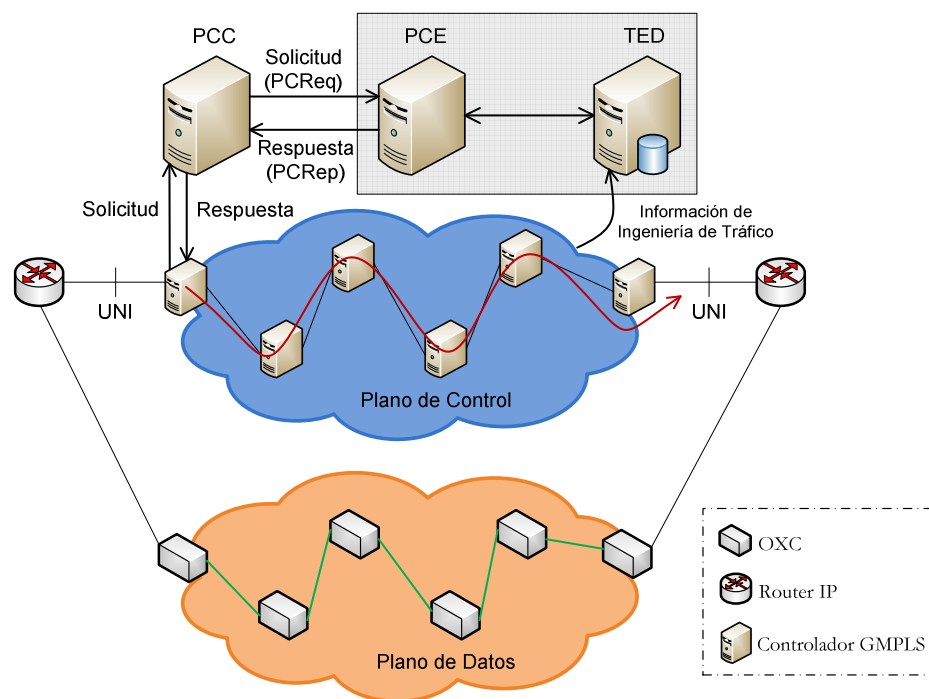


Figura 2.2 - PCE en el Plano de Control

Existen otras redes donde no hay un plano de control y se gestionan desde el Sistema de Gestión de Red (Network Management System - NMS). En estos casos el PCE forma parte del NMS (Figura 2.3) dentro del plano de gestión, de manera que cuando el cliente envía una solicitud de servicio, el NMS solicita un camino al PCE. Dicho PCE necesita la información de la red contenida en la TED para calcular el camino y, una vez que lo ha calculado, envía la respuesta al NMS. Desde el NMS se envía la configuración a los elementos de red para establecer el camino [9].

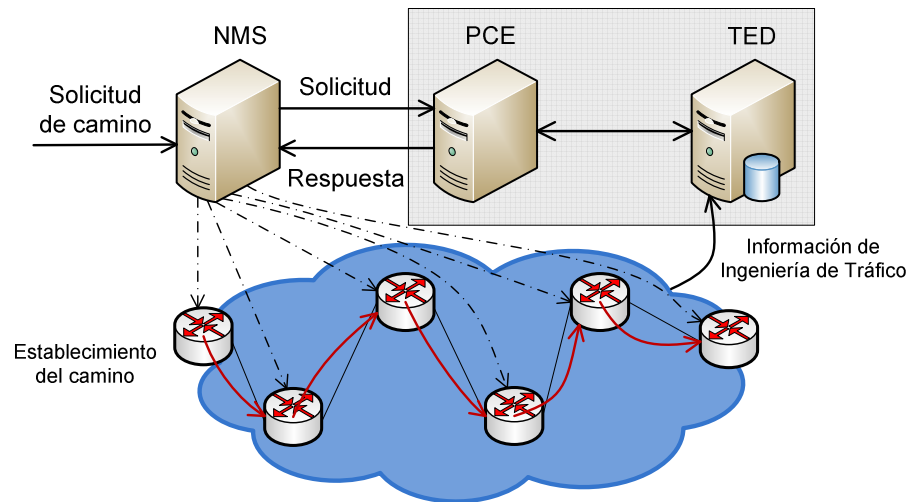


Figura 2.3 - PCE en el Plano de Gestión

El PCE puede estar localizado junto con el PCC o como un servidor independiente (Figura 2.4). La solución conjunta es más fácil de implementar y no requiere de un protocolo de solicitud/respuesta. En cambio, la solución de un servidor separado utiliza el protocolo estándar de solicitud/respuesta PCEP y permite que un PCE dé servicio a un conjunto de PCCs [8].

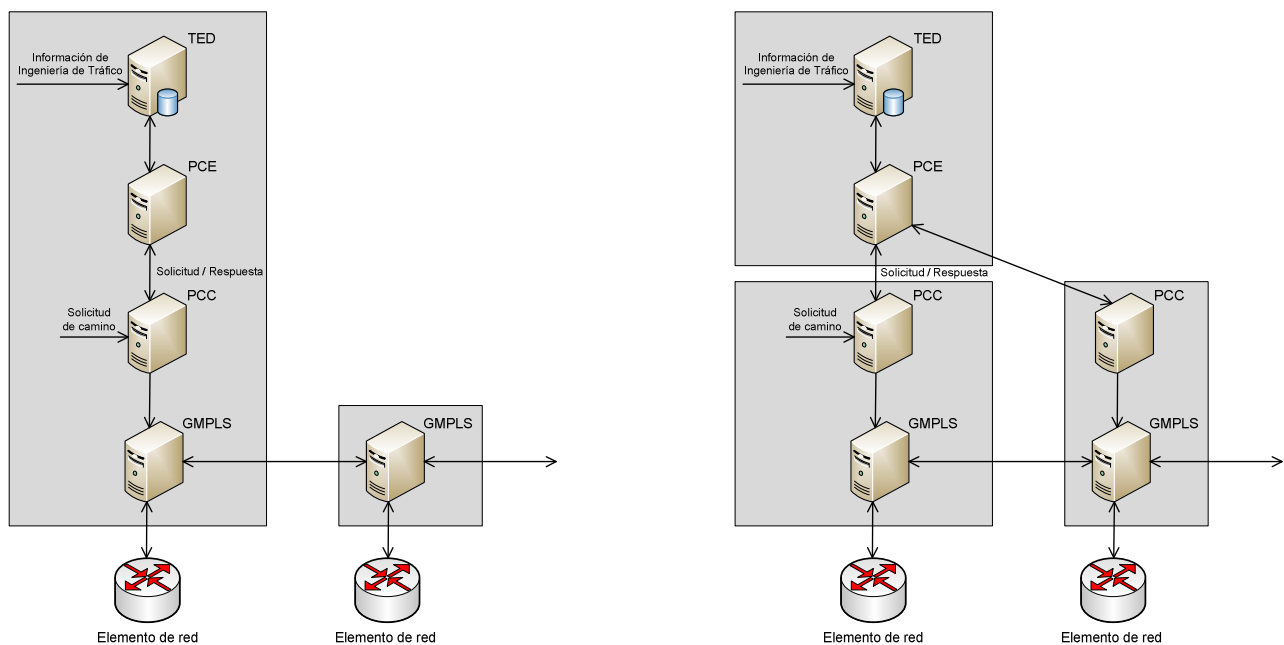


Figura 2.4 - Configuraciones conjunta (izda.) y separada (dcha.) del PCE

Descripción funcional

La Figura 2.5 muestra una secuencia típica en el proceso de cálculo de un TE LSP:

1. Una solicitud de un nuevo camino llega a un dominio MPLS/GMPLS, por lo que el Label Edge Router (LER) de entrada debe empezar el establecimiento de un camino con diferentes restricciones hacia el destino.
2. El LER actúa como un PCC y solicita al PCE el cálculo del mejor camino posible dentro de su dominio. En la comunicación utilizan PCEP.
3. El PCE puede calcular el camino de manera individual o colaborando con otros PCEs del mismo dominio. Para ello comprueba si las restricciones solicitadas pueden ser satisfechas basándose en la información obtenida de la TED y en sus políticas locales.
4. El PCE calcula el camino mediante un algoritmo de cómputo.
5. Cuando el camino es calculado, el PCE se lo proporciona al PCC y éste establece dicho camino mediante el protocolo de señalización utilizado.

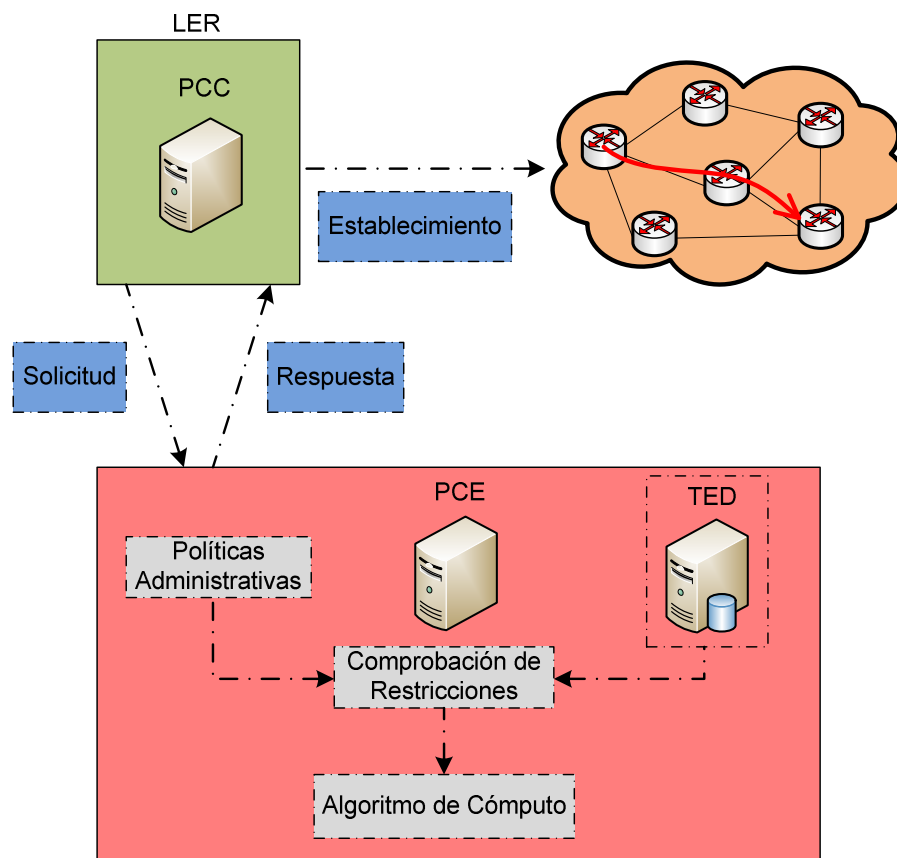


Figura 2.5 - Descripción funcional de la arquitectura basada en PCE

Una arquitectura basada en PCE funciona mediante un modelo computacional, que puede ser *centralizado* o *distribuido*. En un modelo *centralizado* (Figura 2.6), un único PCE se encarga de atender las peticiones de todos los PCCs de su dominio. Esto conlleva el riesgo de tener un punto único de fallos, por lo que una buena opción es designar un segundo PCE (backup) por si el primero falla. Los dos han de tener las mismas políticas locales, pero sólo uno de ellos estará activo en cada dominio [8].

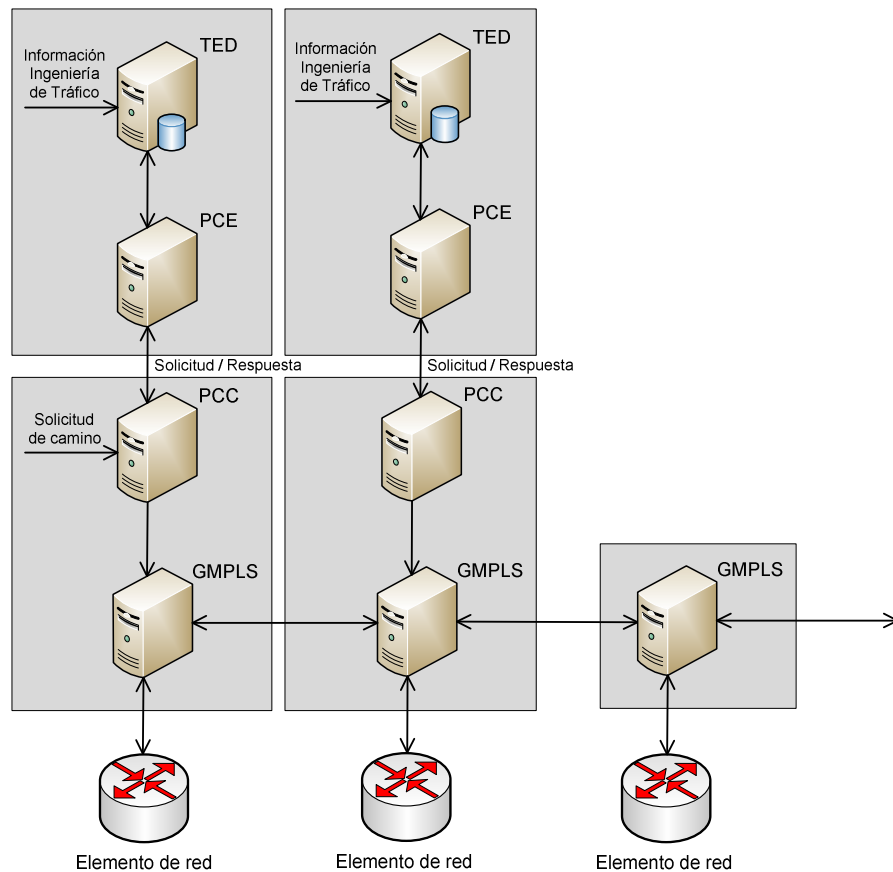


Figura 2.6 - Modelo *centralizado* de la arquitectura PCE

En cambio, en un modelo *distribuido* (Figura 2.7) varios PCEs pueden participar en el cálculo de las rutas ya sea calculando segmentos de ruta o bien comunicándose entre ellos. Un PCC puede estar vinculado a un PCE en concreto o bien elegir el más adecuado libremente. El proceso de elección entre los distintos PCEs se detalla en [10]. Según este modelo, un PCC realiza una solicitud a un PCE, pero éste es incapaz de devolver un camino completo y el siguiente elemento de red puede considerar necesario completarlo o mejorarlo. Esto puede ocurrir cuando el camino devuelto es simplemente un segmento y no alcanza el destino solicitado, o bien cuando se trata de un camino no estricto (loose path). Por tanto, el siguiente elemento de red consulta a otro PCE para establecer el próximo segmento del camino [8].

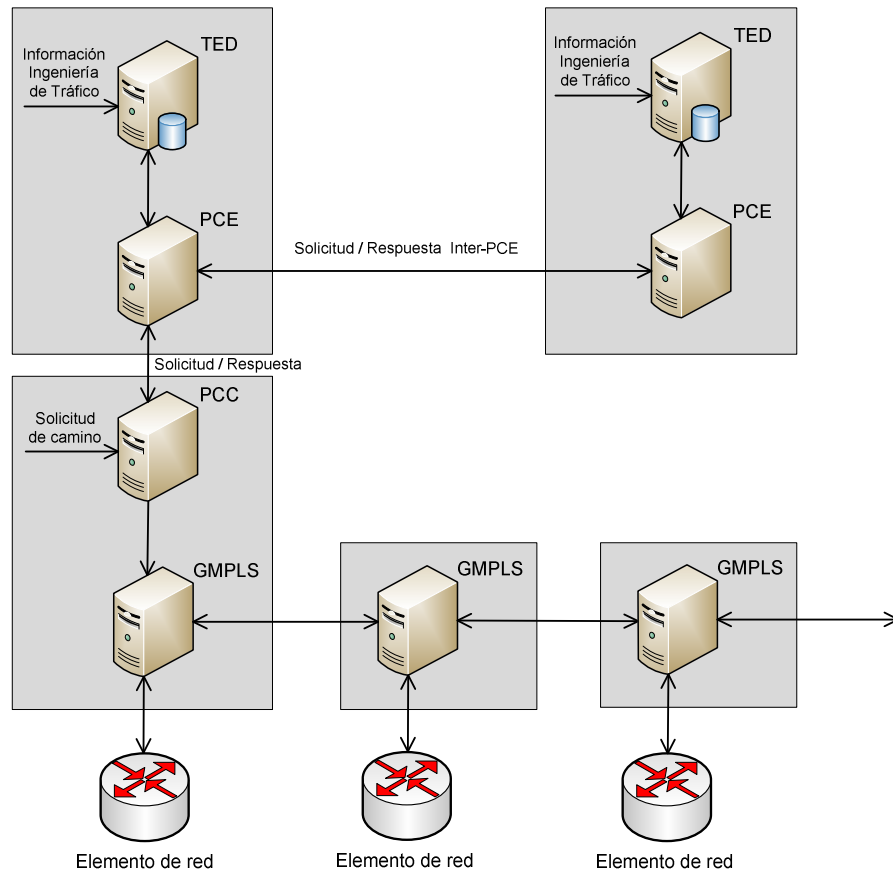


Figura 2.7 - Modelo *distribuido* de la arquitectura PCE

El problema anterior puede resolverse introduciendo comunicación entre los PCEs, de manera que cuando un PCC solicita un camino a un PCE cualquiera, éste puede consultar a otros PCEs para devolver un camino óptimo. Este modelo no proporciona realmente un algoritmo distribuido de cómputo, pero permite que distintos PCEs formen parte del cálculo de un mismo camino [8].

Debe destacarse que un PCC no distingue entre un modelo centralizado o uno con múltiples PCEs con comunicación inter-PCE. El PCC envía una solicitud de cálculo de camino y recibe un camino completo o bien un segmento del mismo, sin ser consciente de si más de un PCE ha colaborado en la consecución del mismo [8].

A menudo determinados servicios requieren del cálculo de varios caminos por diferentes motivos, entre los que se encuentran, por ejemplo, la protección (rutas alternativas) o el balanceo de carga. En tal caso, un PCC puede optar por dos soluciones posibles. La primera de ellas es enviar a uno o varios PCEs múltiples solicitudes de manera individual. En este caso las solicitudes no están sincronizadas. La segunda solución se basa en enviar una única solicitud a un PCE preguntando por el cálculo de un conjunto de caminos de manera sincronizada, pero con la opción de indicar que no es necesaria la sincronización. Normalmente un cálculo sincronizado de los caminos da mejor resultado, pero aumenta el tiempo de cómputo [8].

¿Dónde se encuentra el PCE?

Como ya se ha indicado antes, un PCC puede estar vinculado a un PCE o elegir el más adecuado en función de las restricciones del camino o si se necesita balancear la carga. Para ello debe saber dónde están situados los PCEs de su dominio y qué capacidad de cómputo tiene cada uno. Esta información puede ser configurada manualmente o descubierta de manera dinámica.

Las extensiones de los protocolos IGP (Interior Gateway Protocol) tales como Open Shortest Path First (OSPF) [11] o Intermediate System–Intermediate System (IS-IS) [12] permiten dar a conocer esta información a los PCCs junto con otra adicional referente al estado de los PCEs: detección de un nuevo PCE, un PCE ha fallado, información modificada, etc. [11]. Puede encontrarse más información acerca de dichas extensiones en [13] y [14].

La información de estado de la red es enviada en objetos TLV (Type-Length-Value). Más concretamente, se ha definido un objeto denominado *PCE Discovery TLV*, que contiene información acerca de la dirección IP del PCE, el dominio en el que se encuentra, el/los dominios donde está operativo o sus dominios vecinos [11] y [12].

2.1.2. Path Computation Element Communication Protocol (PCEP)

PCEP es un protocolo basado en el modelo de solicitud/respuesta y utilizado para las comunicaciones entre un PCC y un PCE o entre dos PCEs [15].

PCEP opera sobre el protocolo de transporte TCP mediante sesiones cliente-servidor. Para ello, utiliza siete tipos de mensajes: *Open*, *Keepalive*, *Request*, *Reply*, *Notification*, *Error* y *Close* [15]. Un PCC puede mantener sesiones PCEP con múltiples PCEs y, de manera similar, un PCE puede hacerlo con más de un PCC. En cambio, sólo se podrá mantener una única sesión a la vez entre los mismos pares de la comunicación (PCC-PCE o PCE-PCE).

A continuación se resumen las distintas fases de la comunicación mediante el protocolo PCEP [15] (véase la Figura 2.8):

- **Fase de Inicialización:** se distinguen dos pasos dentro de esta fase. El primero es el establecimiento de una conexión TCP entre un PCC y un PCE y el segundo, el establecimiento de una sesión PCEP sobre la conexión TCP. Una vez establecida una conexión TCP, los dos pares de la comunicación negocian distintos parámetros para establecer una sesión PCEP. Dichos parámetros se envían en mensajes *Open* e incluyen los temporizadores *Keepalive* y *DeadTimer*, además de otra información adicional que determina las condiciones en que un mensaje *Request* ha de ser enviado a un PCE. Para reconocer los mensajes *Open* se utilizan los mensajes *Keepalive*.
- **Sesión Keepalive:** cuando hay una sesión establecida, tanto el PCC como el PCE tienen la oportunidad de conocer si el otro extremo de la sesión sigue estando disponible. Para ello utilizan un mecanismo basado en mensajes *Keepalive* y temporizadores *Keepalive* y *DeadTimer*. Cada extremo de la conexión utiliza un temporizador *Keepalive* que se reinicia tras el envío de cualquier mensaje PCEP. Cuando el temporizador expira se envía un mensaje *Keepalive*. A su vez, los dos extremos también reinician un temporizador *DeadTimer* cada vez que reciben un mensaje PCEP, de manera que si no reciben ningún mensaje antes de que el *DeadTimer* expire, entonces declaran muerta la sesión. Los valores de los temporizadores son indicados en los mensajes *Open*. Un extremo especifica un valor para el temporizador *Keepalive* y recomienda otro para el *DeadTimer* del otro

extremo. El valor mínimo del temporizador *Keepalive* es de 1 segundo y se recomienda un valor de 30 segundos. También existe la opción de deshabilitarlo poniéndolo a cero. En cuanto al *DeadTimer*, su valor recomendado es 4 veces el valor del temporizador *Keepalive* del otro extremo, lo que evita congestionar la red con mensajes *Keepalive*.

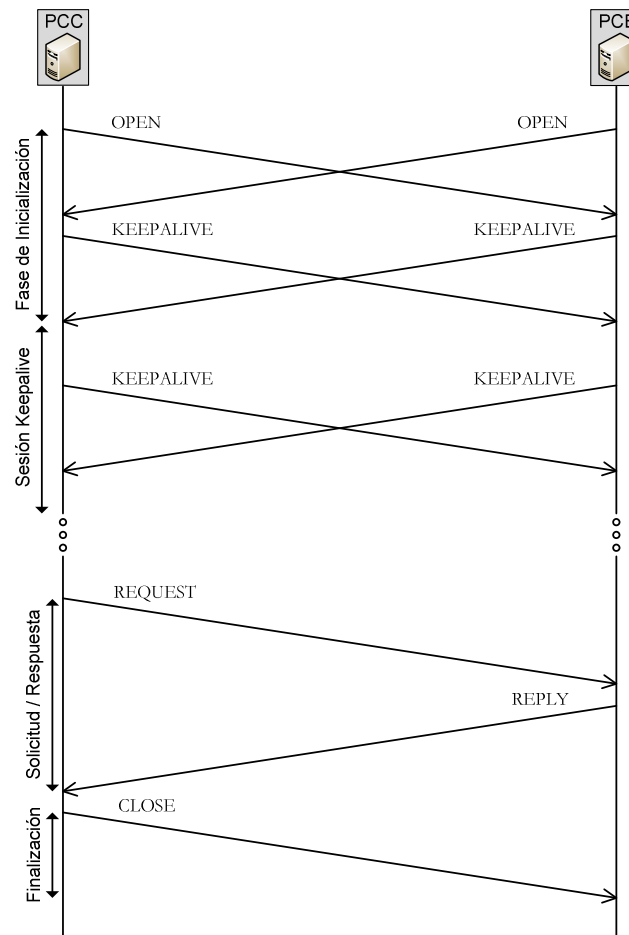


Figura 2.8 - Sesiones PCEP

- Path Computation Request:** cuando se ha establecido una sesión y un PCC ha seleccionado un PCE entre los de su dominio, dicho PCC puede solicitar rutas al PCE utilizando mensajes *Request*. Un mensaje *Request* contiene varios objetos que especifican las restricciones y atributos para el cálculo del camino y está identificado unívocamente por un identificador de petición. Entre la información que contiene el mensaje *Request* se encuentran, por ejemplo, las direcciones IP del origen y el destino o, si la ruta solicitada es multi-dominio, puede indicarse el procedimiento a seguir. Un PCC puede solicitar un conjunto de rutas en el mismo mensaje *Request*, así como un PCE puede proporcionar varias respuestas para un mismo camino.
- Path Computation Reply:** después de haber recibido una solicitud de ruta, el PCE realiza el cálculo correspondiente, cuyo resultado puede ser positivo o negativo. En el caso de se haya podido obtener un camino que satisfaga el conjunto de restricciones solicitadas, el PCE envía el camino o el conjunto de caminos al PCC en un mensaje *Reply*. En cambio, si no se ha podido

encontrar ningún camino, el PCE se lo comunica al PCC indicando, si lo considera necesario, las restricciones que no se han podido satisfacer.

- **Finalización de la sesión PCEP:** cuando uno de los dos extremos desea terminar la sesión PCEP, debe mandar un mensaje *Close* al otro extremo y después cerrar la conexión TCP. Si es el PCE el que finaliza la sesión, el PCC deberá borrar todos los estados relacionados con las peticiones pendientes de ser atendidas. De forma similar, si el PCC es el que cierra la sesión, el PCE borrará todas las peticiones de ese PCC pendientes de ser calculadas. Debe tenerse en cuenta que como PCEP opera sobre TCP, si la conexión TCP falla, la sesión PCEP se cerrará inmediatamente.

Una sesión PCEP puede ser intermitente o permanente. En el modo intermitente, cada vez que se solicita un camino se debe abrir una nueva sesión PCEP. En cambio, en el modo permanente la sesión PCEP permanece activa indefinidamente, lo cual puede ser muy útil cuando el envío de peticiones se realiza con gran frecuencia.

Notificaciones

Es posible que el PCE necesite notificar algún evento a un PCC (o viceversa) mediante un mensaje *Notification*. Por ejemplo (véase la Figura 2.9), el PCE puede informar acerca de la sobrecarga de mensajes que sufre, así como un PCC puede cancelar peticiones si lo considera necesario.

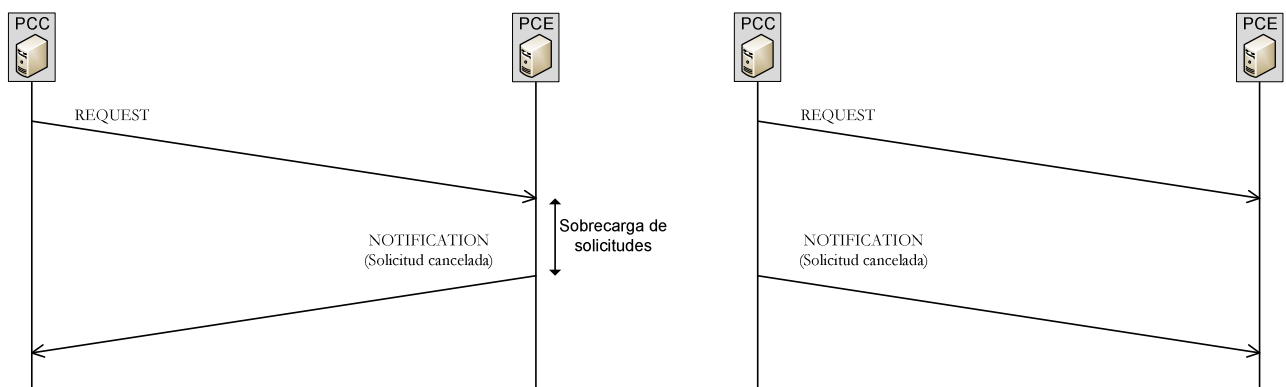


Figura 2.9 – Notificaciones enviadas por el PCE (izda.) y el PCC (dcha.)

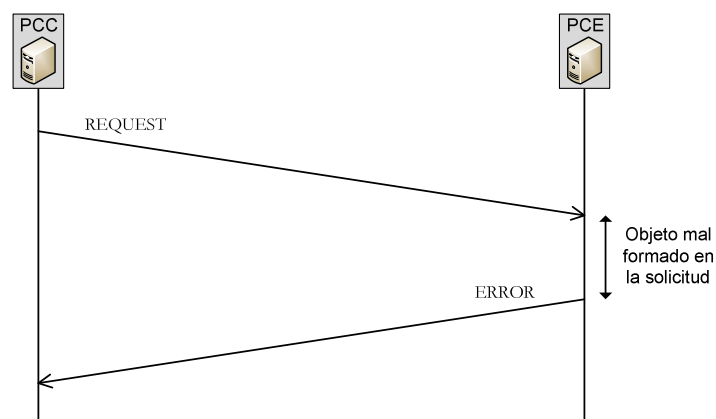
En la Tabla 2.1 se muestra la lista de notificaciones del protocolo PCEP, identificadas por un tipo y un valor de notificación [15].

Tabla 2.1 - Tipos de Notificaciones del protocolo PCEP

Tipo de Notificación	Valor de Notificación
1. Solicitud pendiente cancelada	1. Un PCC cancela un conjunto de solicitudes pendientes debido a motivos tales como la recepción de una respuesta positiva desde otro PCE o un fallo en la red.
	2. Un PCE cancela un conjunto de solicitudes pendientes.
2. PCE sobrecargado	1. Un PCE indica a un PCC que se encuentra sobrecargado. Opcionalmente puede indicar el tiempo estimado que durará la congestión.
	2. Un PCE indica a un PCC que ha salido del estado de sobrecarga en que se encontraba y ya se encuentra disponible.

Errores

Un mensaje de *Error* puede ser enviado en diferentes situaciones que incluyen, por ejemplo, la violación de las especificaciones del protocolo o el alcance de alguna condición de error. En la Figura 2.10 se observa cómo un PCE envía un mensaje de *Error* porque ha recibido un objeto mal formado en el mensaje *Request*.

Figura 2.10 – Mensaje de *Error* enviado por el PCE

En la Tabla 2.2 se proporcionan los posibles errores del protocolo PCEP, identificados por un tipo y un valor de error [15].

Tabla 2.2 – Tipos de Errores del protocolo PCEP

Tipo de Error	Valor de Error
1. Fallo en el establecimiento de la sesión PCEP	1. Recepción de un mensaje <i>Open</i> no válido.
	2. No se recibe un mensaje <i>Open</i> antes de que expire el temporizador <i>OpenWait</i> .
	3. Características de la sesión PCEP inaceptables e innegociables.
	4. Características de la sesión PCEP inaceptables pero negociables.
	5. Recepción de un segundo mensaje <i>Open</i> con características inaceptables durante el establecimiento de la primera sesión.
	6. Recepción de un mensaje de <i>Error</i> en respuesta a un mensaje <i>Open</i> con características inaceptables.
	7. No se recibe un mensaje <i>Keepalive</i> o un mensaje de <i>Error</i> antes de que expire el temporizador <i>KeepWait</i> .
2. Capacidad no soportada	---
3 Objeto desconocido	1. Clase de objeto no reconocida.
	2. Tipo de objeto no reconocido.
4 Objeto no soportado	1. Clase de objeto no soportada.
	2. Tipo de objeto no soportado.
5 Violación de política	1. La bandera <i>C</i> del objeto <i>METRIC</i> en el mensaje <i>Request</i> está activa y no se acepta.
	2. La bandera <i>O</i> del mensaje <i>Request</i> está activa y no se acepta.
6. Falta un objeto obligatorio	1. Falta el objeto <i>RP</i> .
	2. Falta el objeto <i>RRO</i> en una solicitud de reoptimización.
	3. Falta el objeto <i>END-POINT</i> .
7. Falta alguna solicitud sincronizada	---
8. Identificador de solicitud desconocido	---
9. Intento de establecimiento de otra sesión PCEP	---
10. Recepción de un objeto no válido	1. Recepción de un objeto con la bandera <i>P</i> no activa, en contra de las especificaciones impuestas.

2.2. Entornos Multi-dominio

Hoy en día, las redes de comunicaciones están formadas por otras sub-redes más pequeñas denominadas dominios o Sistemas Autónomos (Autonomous System - AS). Un dominio se define como un conjunto de elementos de red administrados bajo la dirección de una sola entidad y que cuentan con una misma política de enrutamiento. Internet, por ejemplo, es el resultado de la interconexión de casi 38.000 ASes [16].

Dichos dominios o ASes pueden ser definidos en base a razones de propiedad, a motivos administrativos o de geografía. En la Figura 2.11 se muestra un escenario multi-dominio con cinco dominios. En cada dominio, los nodos que conectan con otros dominios se denominan nodos de borde. El resto son nodos interiores que están conectados con el resto de nodos del dominio. Todos los operadores administran un único AS, excepto el operador A, que administra dos.

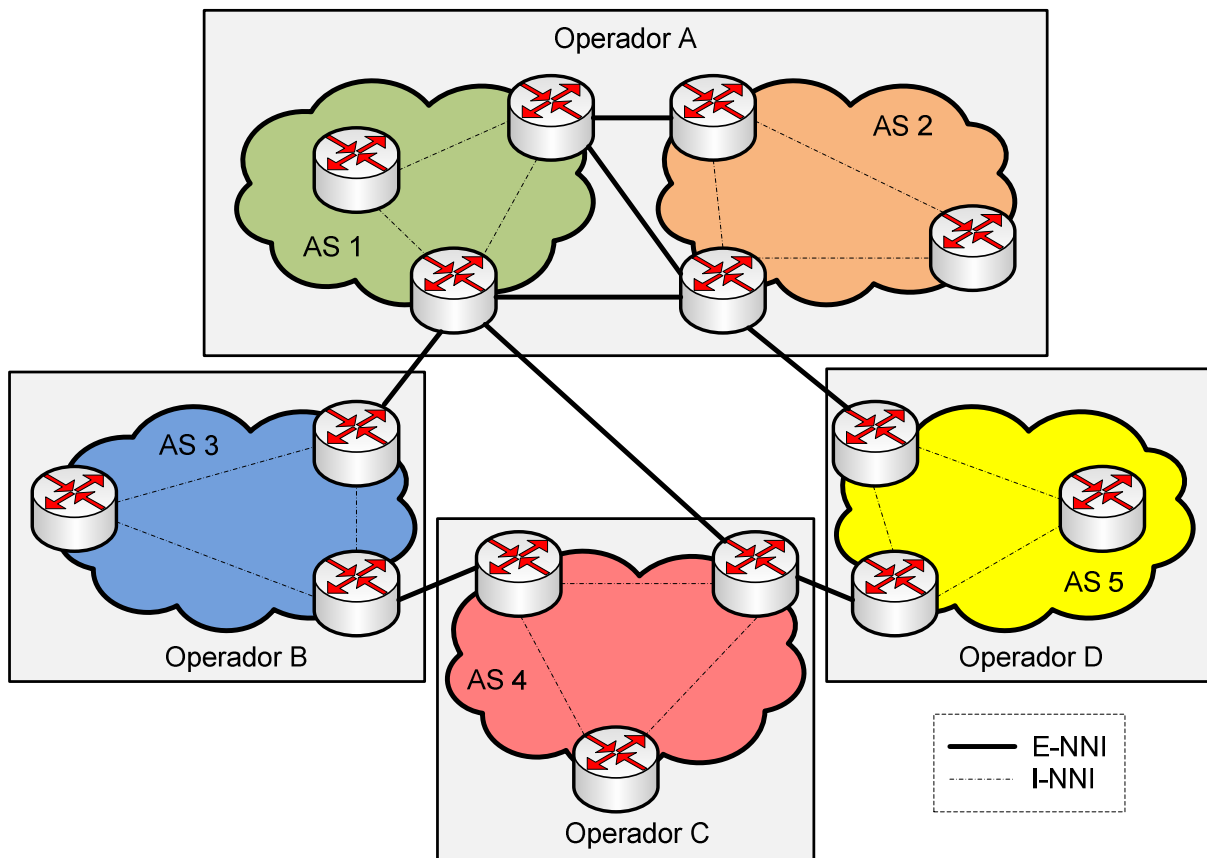


Figura 2.11 – Escenario con múltiples dominios o ASes de diferentes operadores

Para administrar las comunicaciones entre los distintos dominios, es necesario el intercambio de información inter-dominio e intra-dominio por las Interfaces Red a Red (Network-Network Interface - NNI). De acuerdo a la nomenclatura ASON [1], la interfaz inter-dominio se denomina Interfaz Externa de Red a Red (External Network-Network Interface - E-NNI) y la intra-dominio Interfaz Interna de Red a Red (Internal Network-Network Interface - I-NNI). Dichas interfaces deben soportar, al menos, dos funcionalidades: el enrutamiento y la señalización.

Las soluciones existentes para el enrutamiento y la señalización están limitadas para su funcionamiento dentro del AS o área de enrutamiento del nodo de entrada del LSP. Esto es así porque dicho nodo de entrada tiene visibilidad limitada de la topología y los recursos de ingeniería de tráfico a su propio dominio y no puede calcular un camino óptimo que recorra múltiples dominios. Sin embargo, un camino inter-dominio óptimo debería ser entendido como el camino más corto con restricciones que podría ser calculado como si sólo hubiera un único dominio de enrutamiento y la información referente a la topología y los recursos de ingeniería de tráfico estuviera disponible en su totalidad [16]. Partiendo de esta premisa, a continuación se resumirán las técnicas actuales de enrutamiento y señalización inter-dominio.

2.2.1. Enrutamiento en redes multi-dominio

El BGP (Border Gateway Protocol) [18] es un protocolo mediante el cual se produce el intercambio de información de enrutamiento entre diferentes ASes. Los protocolos de enrutamiento intra-dominio, tales como OSPF [26] o IS-IS [25], reúnen la información de su dominio y la envían a los routers de borde, los cuales intercambian dicha información utilizando BGP. Existen dos escenarios multi-dominio posibles de configuración del conjunto de direcciones IP: separado (Figura 2.12) o compartido (Figura 2.13).

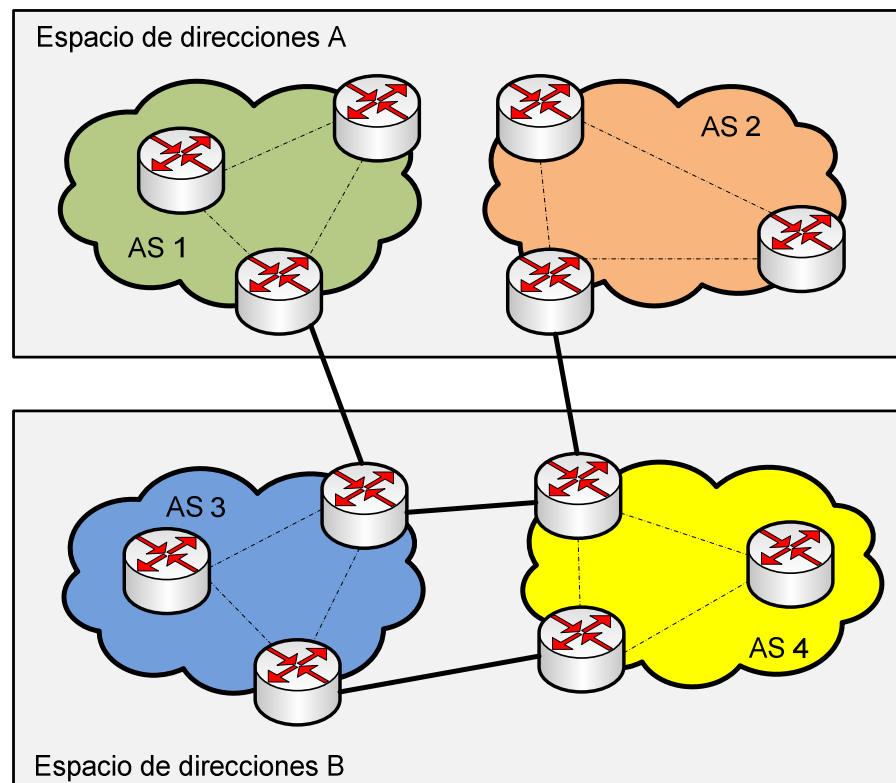


Figura 2.12 - Espacios de direcciones IP separados

En la Figura 2.12 se pueden ver dos espacios de direcciones distintos (A y B). El AS 1 y el AS 2 comparten el espacio de direcciones A y el AS 3 comparte con el AS 4 el espacio de direcciones B. La estructura superpuesta de este modelo se asemeja a la arquitectura ASON. Además, puede ser útil cuando

los dominios del espacio de direcciones B son redes ópticas que proporcionan conectividad a los dominios del espacio de direcciones A. En la Figura 2.13 el espacio de direcciones es compartido.

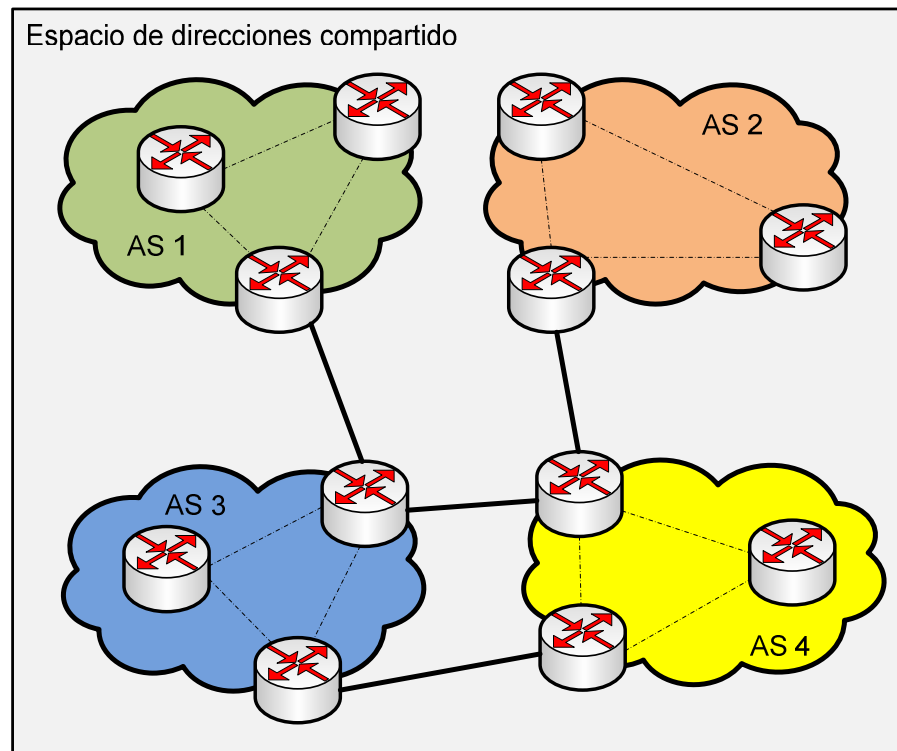


Figura 2.13 - Espacio de direcciones IP compartido

A pesar de que BGP ha dominado durante décadas el campo del enrutamiento inter-dominio, no soporta el enrutamiento con extensiones de ingeniería de tráfico o calidad de servicio por motivos de escalabilidad y confidencialidad. Esto quiere decir que las redes que utilizan BGP sólo intercambian información referente a la conectividad, lo que es insuficiente para satisfacer ciertos requisitos en el cálculo de TE LSPs inter-dominio y optimizar la utilización de la red. Este problema ha sido parcialmente resuelto por BGP en [19].

2.2.2. Señalización en redes multi-dominio

La Figura 2.14 muestra una red de transporte multi-dominio basada en la arquitectura ASON. Los clientes están conectados a la red de transporte mediante la Interfaz Usuario-Red (User-Network Interface - UNI) [20]. Cada AS tiene su propio protocolo de señalización. Por lo tanto, como se muestra en la Figura 2.14, UNIs, E-NNIs y I-NNIs mantienen distintas sesiones del protocolo RSVP (Resource reSerVation Protocol) [7], concretamente cinco. Sin embargo, la inclusión de GMPLS en todos los dominios de la red hace posible establecer una única sesión RSVP en todas las interfaces de red. Los autores de [19] han llevado a cabo pruebas con el objetivo de observar la interoperabilidad entre ASON y GMPLS.

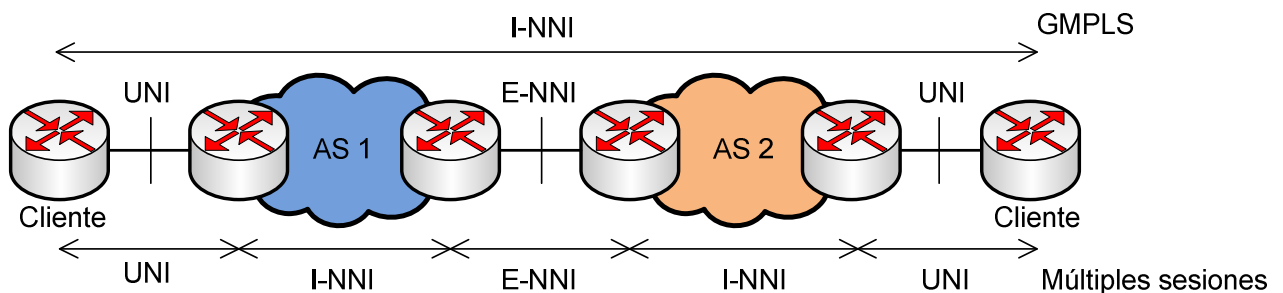


Figura 2.14 - Señalización en una red multi-dominio

2.2.3. Problema del cálculo de caminos multi-dominio

Con el fin de ilustrar el problema del cálculo de caminos multi-dominio, se puede observar el escenario de la Figura 2.15 en donde el nodo A del dominio 1 necesita establecer un camino hasta el nodo B del dominio 3. Como el nodo A sólo tiene visibilidad local (dentro de su dominio), el cálculo del camino podría ser distribuido por los nodos de borde intermedios, lo que supondría un coste elevado, o bien podría ser responsabilidad de los distintos PCEs a lo largo de la secuencia de dominios.

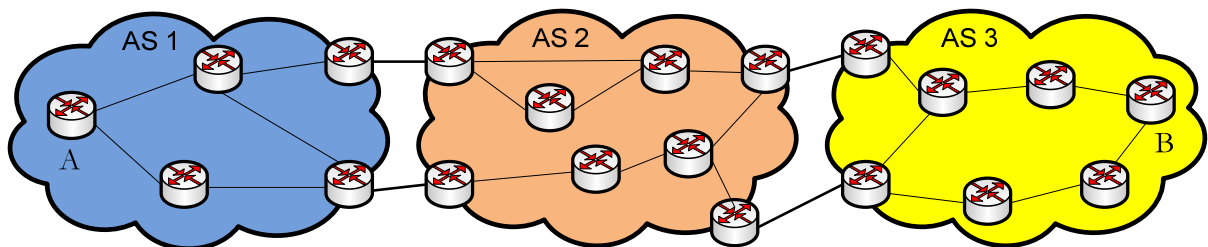


Figura 2.15 - Escenario multi-dominio

En este proyecto se han estudiado distintas soluciones para el cálculo de caminos multi-dominio basadas en una arquitectura PCE: *Per-domain Path Computation*, *Simple Cooperating PCEs*, *Backward-Recursive PCE-based Computation* y *PCE Jerárquico*. Estas soluciones se explican en el tercer capítulo.

2.2.4. Tipos de bloqueo para PCE en entornos multi-dominio

Una vez presentado el problema del cálculo de caminos multi-dominio y propuesta la solución mediante la utilización de una arquitectura basada en PCE, es conveniente repasar los tipos de bloqueo que se pueden presentar en una red óptica multi-dominio. Un PCE necesita la información referente a las características y estado de los enlaces almacenada en su TED, ya que es la entidad encargada de realizar las tareas de enrutamiento, validación de restricciones físicas y elección de recursos (longitudes de onda) para los caminos. Por tanto, se pueden dar tres tipos de bloqueo en consecuencia de estas tareas.

- El primero de ellos se da cuando no se puede enrutar un camino desde el origen hasta el destino. Esto puede ocurrir porque no existen enlaces para alcanzar el destino debido a fallos en la red.
- El segundo tipo de bloqueo se produce por restricciones físicas. Tras el enrutamiento, el camino debe cumplir con las restricciones físicas de los enlaces. En caso negativo, será rechazado.
- El último tipo ocurre cuando no hay recursos (longitudes de onda) disponibles para el camino, es decir, todas las longitudes de onda de alguno de los enlaces del camino están asignadas a otros caminos, por lo que el solicitante deberá esperar a que se liberen los recursos necesarios.

La información recogida por la TED es actualizada en el plano de control mediante mensajes LSA (Link State Advertisement) del protocolo OSPF [9]. Cuando la información que contiene la TED es diferente del estado real de la red, el PCE puede asignar recursos que ya estaban reservados por otros caminos. Existe también un retardo desde que el PCE envía su respuesta hasta que el LSR establece el camino. Dicho retardo incluye, además del producido en el plano de control, el del intercambio de mensajes PCEP. Cuando el LSR intenta reservar el camino, los recursos ya estaban reservados por otro LSR, por lo que su petición es denegada. Se produce pues un tipo de bloqueo adicional.

Se ha propuesto recientemente un draft en el IETF [35] con extensiones del protocolo PCEP con el objetivo de evitar este último tipo de bloqueo. Lo que se propone es que el PCE pueda pre-reservar durante un tiempo determinado los recursos asignados a una petición de camino de un LSR, por lo que se evita que esos mismos recursos sean utilizados por otros LSRs.

A continuación se exponen los requisitos para el protocolo PCEP propuestos en [35]. Un PCC que solicita el cálculo de un camino, debería ser capaz de indicar en el mensaje *Request*:

- El tiempo que los recursos deberán estar reservados.
- El tipo de recursos que serán reservados (ancho de banda, ranuras de tiempo, longitud de onda, etc.) o la granularidad de los recursos, que se refiere a la posibilidad de reservar no sólo los recursos, sino también los enlaces o nodos involucrados.
- Si los recursos deben ser reservados para posteriores peticiones.

El PCE debería ser capaz de:

- Calcular el o los caminos solicitados y, de acuerdo a las indicaciones del PCC, reservar los recursos elegidos por un periodo de tiempo.
- El tiempo determinado por el PCE nunca debe ser menor que el solicitado por el PCC (puede ocurrir que el PCE sólo pueda reservar recursos para unos tiempos en concreto, múltiplos de un valor). Alternativamente, podría enviar un mensaje de *Error* al PCC si el camino no puede ser calculado o pre-reservado.
- También debería ser capaz de aplicar una granularidad diferente a la indicada en la solicitud. En tal caso, debe comunicárselo al PCC.

Además, el PCE debería poder responder en el mensaje *Reply* lo siguiente:

- Si los recursos han sido pre-reservados correctamente y el periodo de tiempo utilizado, que podría ser diferente del solicitado.
- La granularidad de los recursos reservados, que podría ser diferente a la indicada por el PCC.
- Proporcionar, por ejemplo, un identificador de pre-reserva para que el PCC pueda cancelarla.

En base a estos requisitos, se proponen en [35] una serie de extensiones al protocolo PCEP para su correcto funcionamiento. Un PCC que solicita un camino al PCE, debe incluir en su mensaje *Request* dos objetos: *PCC_ID_REQ* y *RESERVATION*. El primero de ellos, cuya información puede encontrarse en [36], se usa para indicar la dirección IP del PCC. El objeto *RESERVATION* [35] indica la intención del PCC de pre-reservar los recursos asignados a su petición. En él puede indicarse el tipo de recurso solicitado y el tiempo que se quiere que dicho recurso sea reservado. El PCE, en respuesta a la petición del PCC y opcionalmente, puede incluir un objeto *RESERVATION_CONF* [35] en el mensaje *Reply* indicando el tiempo que el recurso ha sido reservado finalmente y el identificador de la reserva.

También se propone la posibilidad de que un PCC pueda cancelar una reserva indicándolo en un mensaje *Notification*. Para ello debería incluir el identificador de la reserva en un objeto *RESERVATION_ID TLV* [35].

3. Soluciones basadas en PCE para entornos multi-dominio

El cálculo de caminos inter-dominio representa un problema debido a que, como se ha explicado en el capítulo anterior, no existe un único elemento en la red que tenga acceso a la información de los enlaces y recursos de todos los dominios. Utilizando una arquitectura basada en PCE, existen varios procedimientos que dan solución a este problema y que se estudiarán en este capítulo.

En la sección 3.1 se presenta el método *Simple Cooperating PCEs*, en la sección 3.2 se examina el *Per-domain Path Computation*, en la sección 3.3 el procedimiento *Backward-Recursive PCE-based Computation (BRPC)* y en la sección 3.4 una solución basada en una arquitectura jerárquica de los PCEs: *PCE Jerárquico (H-PCE)*.

3.1. Per-Domain Path Computation

Se trata de un método de establecimiento de caminos inter-dominio [21] en donde el camino es calculado durante el proceso de señalización dominio a dominio. Por tanto, no existe comunicación entre los PCEs de diferentes dominios. El nodo de entrada de cada dominio es el encargado de establecer el segmento del camino que atraviesa su dominio. Para ello, actuará como un PCC y preguntará al PCE por dicho segmento. Cuando haya recibido una respuesta, establecerá la ruta local mediante el protocolo de señalización especificado, hasta llegar al siguiente dominio.

Partiendo del escenario de la Figura 3.1, se puede ver como los PCEs de la secuencia son consultados uno tras otro y los segmentos de camino, establecidos antes de pasar al siguiente dominio.

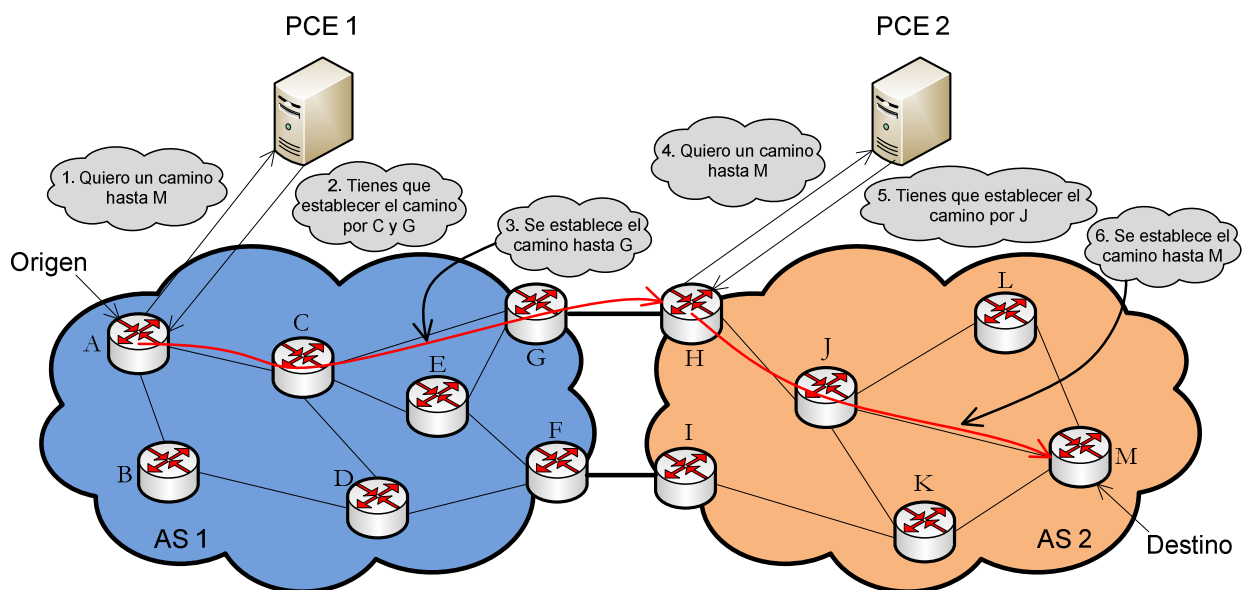


Figura 3.1 – Per-domain Path Computation

Cada cálculo resulta en la mejor ruta local posible que conecte con el siguiente dominio de la cadena (normalmente indicado en el proceso de señalización con un identificador del siguiente dominio o del siguiente nodo de entrada).

Mediante esta técnica es común conseguir caminos de extremo a extremo sub-óptimos, ya que la elección de la mejor ruta local en uno de los dominios condiciona el uso de un nodo de entrada específico en el siguiente dominio de la cadena. En el ejemplo de la Figura 3.1, la elección de G como nodo de salida del primer dominio condiciona al nodo H a actuar de PCC y establecer un camino hasta el destino. Esto puede resultar en una mala ruta local en dicho dominio.

El problema surge a la hora de elegir un nodo de borde de salida, lo cual se soluciona si se conoce a priori la secuencia de dominios que se deben recorrer. En el caso en que dicha secuencia no sea conocida, el PCE debe elegir un nodo de salida basándose en algún criterio de selección para alcanzar el destino. Los autores en [21] sugieren que esto puede lograrse utilizando la ruta IP más corta conocida mediante BGP. Sin embargo, en muchos sistemas GMPLS no se cuenta con enrutamiento IP inter-dominio, por lo que protocolo per-domain podría utilizar la técnica de crankback [22] para conseguir un camino sub-óptimo con el riesgo de que el proceso pueda resultar lento y complicado.

Debe conocerse también que los PCEs de diferentes dominios puede tener distintas características y/o capacidades, utilizar algoritmos de cálculo diferentes o aplicar distintos criterios de optimización. Esto puede resultar en la consecución de un camino inconsistente y sub-óptimo.

El uso de per-domain es conveniente en escenarios en los cuales los dominios están conectados por un sólo enlace y la elección de los nodos de borde que se deben utilizar es fácil de realizar o está predeterminada [23]. En el ejemplo de la Figura 3.1, en donde los dos dominios están interconectados por dos enlaces diferentes, el PCE 1 elige el camino A-C-G-H hacia el destino, siendo el camino global resultante A-C-G-H-J-M, que podría tener, por ejemplo, un coste igual a diez. Sin embargo, podría existir otro camino (A-C-F-I-K-M) con un coste igual a nueve, menor que el anterior. Esto es un problema ya que el PCE 1 no conoce dicha información.

3.2. Simple Cooperating PCEs

En este caso, a diferencia del per-domain, los PCEs de los diferentes dominios pueden intercambiar información para encontrar el mejor camino posible hasta el destino. Cada PCE de la cadena envía al siguiente PCE su mejor ruta local, pero éste puede sugerirle que utilice otra conexión. Véase como ejemplo el escenario de la Figura 3.2, en donde el nodo A (origen) solicita el cálculo de un camino hasta el nodo M (destino), que se encuentra en un dominio diferente. Al recibir la solicitud, el PCE del dominio origen (PCE 1) considera que el camino debe pasar por el nodo de borde F, pero pregunta al siguiente PCE de la cadena (PCE 2), el cual le sugiere que es mejor que el camino pase por G. Por tanto, el PCE 1 responde al PCC con el mejor camino posible desde A hasta G. El nodo de borde del siguiente dominio (H), que está conectado con G mediante un enlace inter-AS, actúa ahora como PCC y solicita al PCE de su dominio (PCE 2) el segmento de camino hasta el destino. Dicho PCE, al encontrarse el destino dentro de su dominio, devuelve el mejor camino posible hasta él. El resultado es un camino global formado por dos rutas locales sub-óptimas.

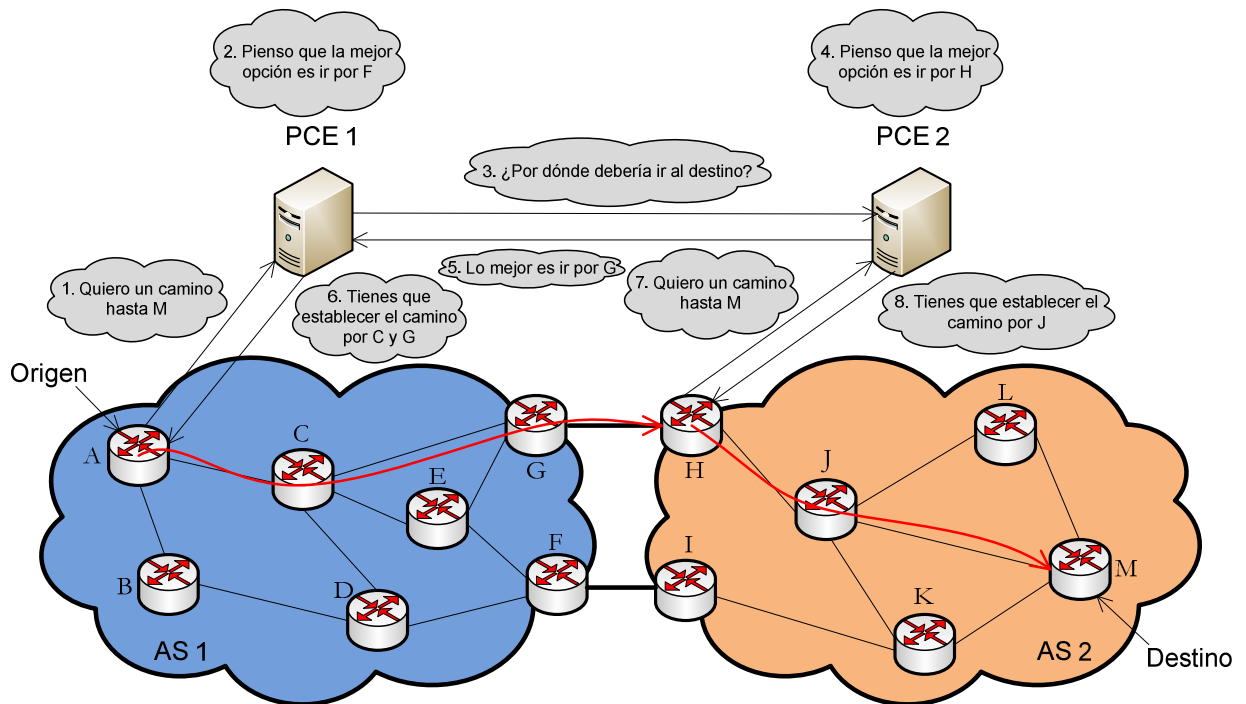


Figura 3.2 – Simple Cooperating PCEs

Las conexiones se eligen localmente, por lo que cuando el camino completo no utiliza las rutas locales óptimas, entonces éste tampoco es óptimo. En ocasiones, incluso, no utiliza ninguna de las mejores rutas locales, como puede observarse en la Figura 3.3. Esto hace patente la falta de optimización a pesar de ser el primer procedimiento en el que la comunicación entre PCEs es posible.

Además, el proceso se complica cuando la secuencia de dominios la forman más de dos, por lo que este procedimiento no se ha implementado en este proyecto.

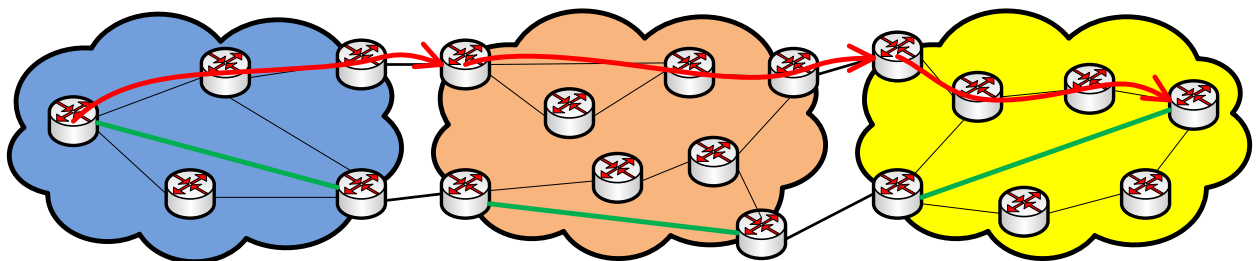


Figura 3.3 – Problema de optimización en Simple Cooperating PCEs

3.3. Backward-Recursive PCE-Based Computation (BRPC)

Este procedimiento [23] se basa en la cooperación y la comunicación entre PCEs para calcular caminos inter-dominio óptimos. El PCC envía una petición al PCE responsable de su dominio, el cual reenvía dicha petición al siguiente PCE de la cadena y así sucesivamente hasta llegar al PCE del dominio destino. Este último PCE crea un árbol denominado Virtual Shortest Path First (VSPT) en el que cada rama representa un camino desde uno de sus nodos de entrada hasta el nodo destino y lo envía al PCE anterior de la cadena en respuesta a su petición. Todos los PCEs de la cadena hasta regresar al dominio origen calculan sus propios caminos desde sus nodos de entrada hasta sus nodos de salida y los añaden al árbol. El PCE del dominio origen debe calcular caminos desde el nodo origen para cada nodo de salida y añadirlos a las ramas del árbol. Posteriormente, deberá elegir el camino óptimo (la rama del árbol con menor coste) y proporcionárselo al PCC que hizo la solicitud.

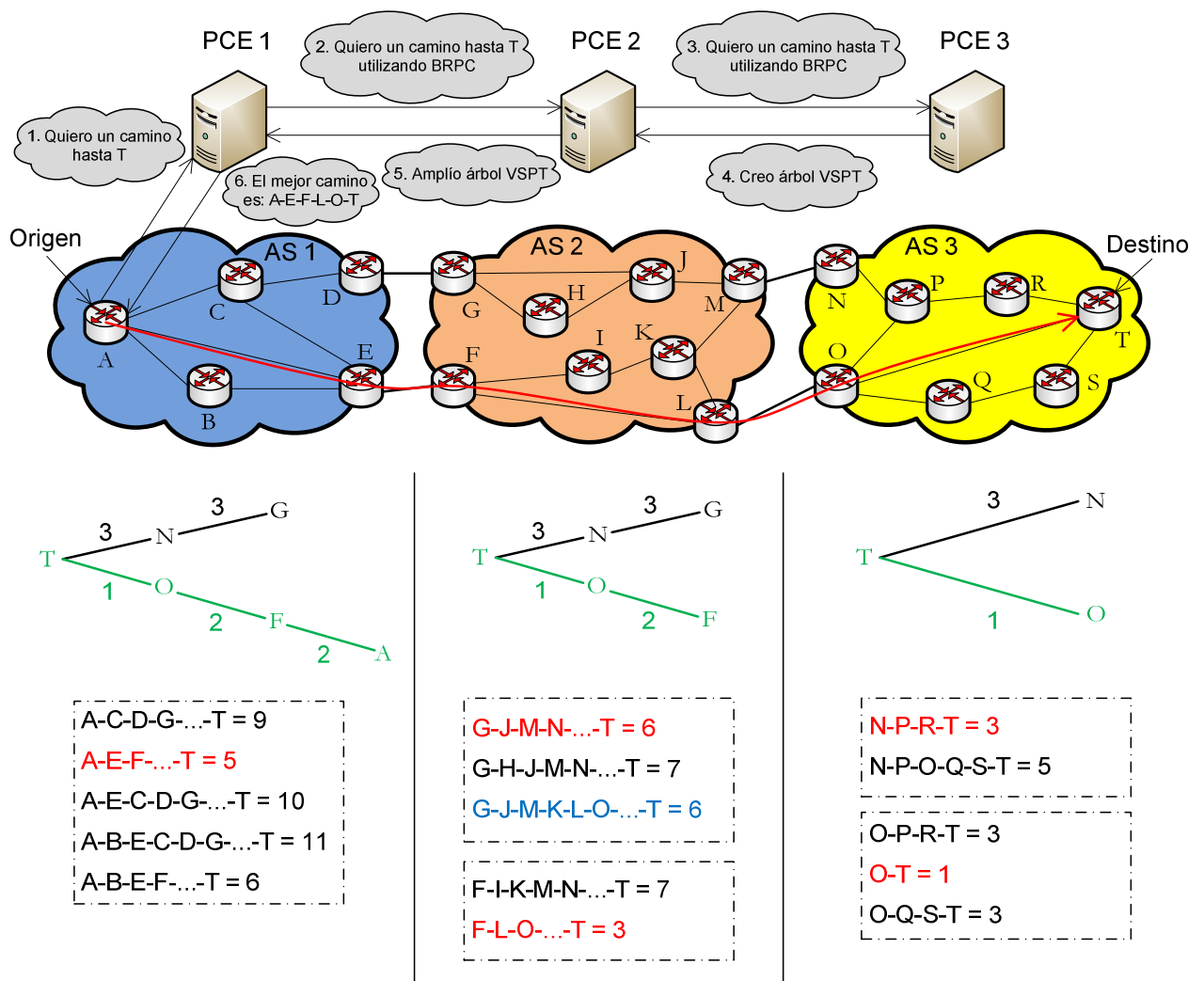


Figura 3.4 – Backward-Recursive PCE-based Computation

En el ejemplo de la Figura 3.4 el nodo A (origen) del dominio 1 necesita establecer un camino hasta el nodo T (destino) del dominio 3, por lo que realiza una solicitud al PCE de su dominio (PCE 1). Dicho PCE, al no poder llevar a cabo el cálculo del camino de manera autónoma, inicia el procedimiento BRPC en cooperación con otros PCEs hasta llegar al destino. El PCE del dominio destino (PCE 3) crea un árbol VSPT cuyas ramas representan sendos caminos desde N y O (nodos de entrada) hasta T (nodo destino). El PCE intermedio (PCE 2), tras recibir el árbol, añade sus propias rutas locales desde F y G (nodos de entrada) hasta L y M (nodos de salida) y lo envía al PCE 1. Éste, con la información contenida en el árbol, es capaz de obtener el mejor camino posible (A-E-F-L-O-T con coste igual a cinco) y proporcionárselo al PCC para su establecimiento.

BRPC se adapta fácilmente a escenarios en donde existen varias conexiones entre los dominios y no hay preferencias sobre cuál de ellas se debe elegir [23]. Se comprueban todas las conexiones con el objetivo de encontrar el mejor camino posible. Un PCE debe conocer los atributos de ingeniería de tráfico de dichas conexiones o enlaces inter-AS, por lo que una solución sencilla consiste en permitir a los nodos de borde inundar esta información a pesar de que no existe adyacencia IGP sobre los enlaces inter-AS. Esto permite obtener al PCE toda la información de ingeniería de tráfico hasta los nodos de borde de los dominios vecinos. Las extensiones de IGP pueden verse en [25] y [26].

Al igual que ocurre con el per-domain, la secuencia de dominios utilizada en BRPC debe ser conocida a priori, por ejemplo, mediante BGP. Un PCC puede indicar dicha secuencia de dominios en un objeto *IRO (Include Route Object)* [15] para que los PCEs sean conscientes de ello. Partiendo de esta premisa, que supone una restricción adicional, el procedimiento BRPC garantiza la consecución de caminos inter-dominio óptimos. Cuando la topología con la que se trabaja está formada por un conjunto de dominios mallados y es posible seleccionar varias secuencias de dominios diferentes, mediante este método se pueden comprobar todas las secuencias y elegir un camino óptimo de entre todas ellas.

Por otra parte, el procedimiento BRPC requiere la especificación de una nueva bandera denominada *VSPT* en el objeto *RP (Request Parameters)* dentro del mensaje *Request* [15]. Cuando esta bandera está activa, significa que el PCC solicita el cálculo de un camino inter-dominio usando el procedimiento BRPC. Por otro lado, dado que los caminos calculados durante todo el procedimiento deben ser proporcionados con sus respectivos costes, la bandera *C* del objeto *METRIC* dentro del mensaje *Request* debe estar activa.

Si el procedimiento no puede completarse porque un PCE de la cadena no reconoce la bandera *VSPT*, dicho PCE envía un mensaje de *Error* al dominio anterior con un tipo de error igual a 4 (*Objeto no soportado*) y un valor igual a 4 (*Parámetro no soportado*). La petición es entonces cancelada. En cambio, si el procedimiento no puede completarse porque un PCE a lo largo de la cadena reconoce, pero no soporta el procedimiento, deberá devolver un mensaje de *Error* con un tipo de error igual a 13 (*El procedimiento BRPC no puede completarse*) y un valor igual a 1 (*Procedimiento BRPC no soportado por uno o más PCEs a lo largo de la cadena*). Recuérdese que los objetos *Error* incluidos en los mensajes *Error* reportados están caracterizados por un tipo de error y un valor, que proporciona información adicional acerca del tipo de error [15].

Si un PCE de la cadena necesita reenviar una solicitud a otro PCE pero no hay ninguno que esté disponible, debe enviar una respuesta negativa al solicitante incluyendo dentro del mensaje *Reply* un objeto *NO-PATH*. Este objeto, a su vez, incluirá un TLV denominado *NO-PATH-VECTOR* [15] con la bandera *Cadena de cálculo BRPC no disponible* activa.

En caso de éxito, el árbol VSPT es enviado dentro del mensaje *Reply* como una lista de objetos *ERO (Explicit Route Object)* en la que cada uno de ellos representa un camino desde el nodo de origen hasta el nodo destino. Por cada nodo del camino se incluye un sub-objeto *ERO* (definido en [6]) especificando la dirección IP del nodo.

El mayor problema de BRPC se da en topologías malladas con múltiples conexiones entre los dominios, en donde el proceso de cálculo de un camino puede volverse una tarea compleja. Si hiciéramos funcionar el procedimiento en el escenario de la Figura 3.5, el PCE 3 proporcionaría árboles VSPT al PCE

2 y al PCE 4, el PCE 2 lo haría con el PCE 4 y viceversa, y el PCE 1 recibiría árboles del PCE 2 y el PCE 4, probablemente varias veces. Además, esto supone una alta carga de mensajes en la red.

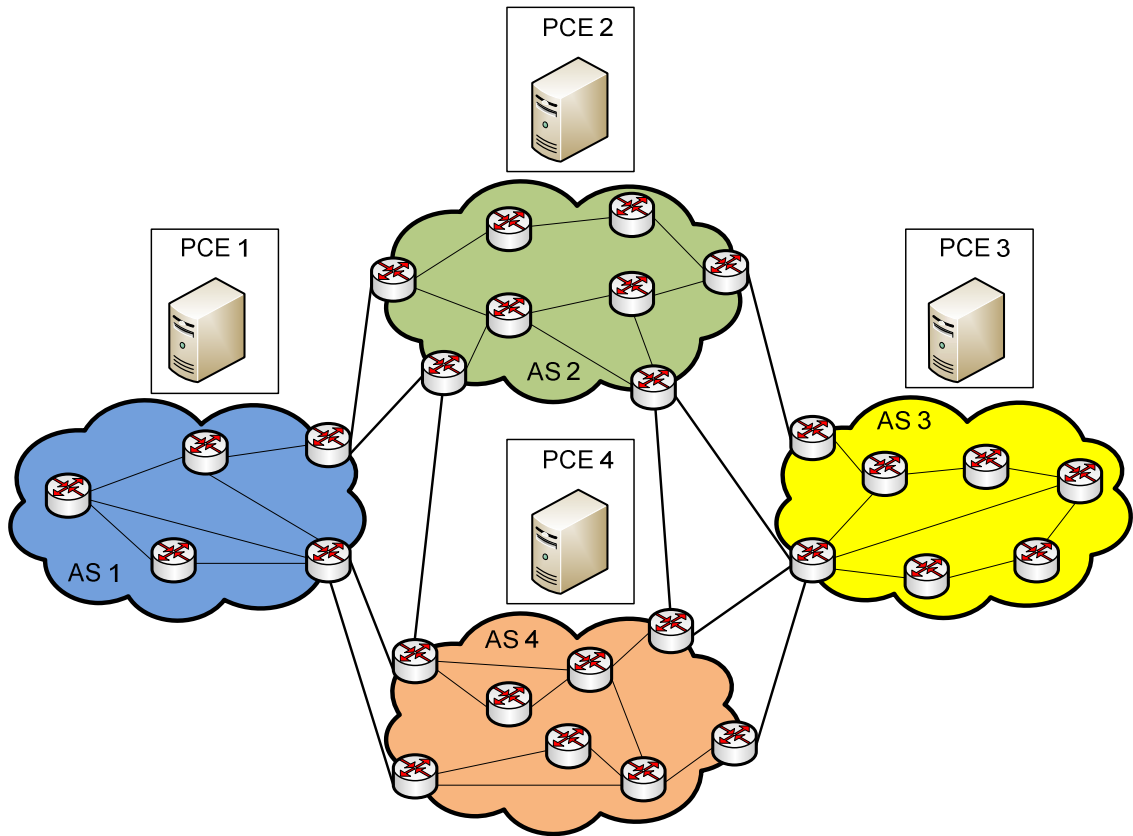


Figura 3.5 – Escenario complejo para Backward-Recursive PCE-based Computation

3.4. PCE Jerárquico (H-PCE)

Como se ha descrito en los apartados anteriores, otras técnicas de cálculo de caminos inter-dominio como per-domain o BRPC asumen de antemano que la secuencia de dominios que se deben atravesar durante el proceso de cálculo es conocida. No se ha dado todavía una solución clara de cómo generar esa secuencia de dominios o qué criterios deben usarse. Es cierto que en escenarios con pocos dominios este proceso de selección no es complicado, pero cuando se trata de otro tipo de topologías (redes ópticas con múltiples sub-dominios o entornos multi-AS) la selección de una secuencia que permita encontrar un camino óptimo desde el origen hasta el destino se vuelve compleja.

Por otro lado, como se ha explicado en el apartado anterior, el uso de BRPC en topologías complejas con múltiples dominios interconectados no es una buena solución.

Es por ello que surge un nuevo método basado en una arquitectura jerárquica de los PCEs (H-PCE) [23], en donde no será necesario conocer a priori la secuencia de dominios a utilizar en el cálculo de caminos. Debe quedar claro que dicho método está orientado para trabajar con pequeños grupos de dominios y no se tiene intención de hacerlo con redes de múltiples dominios como, por ejemplo, Internet.

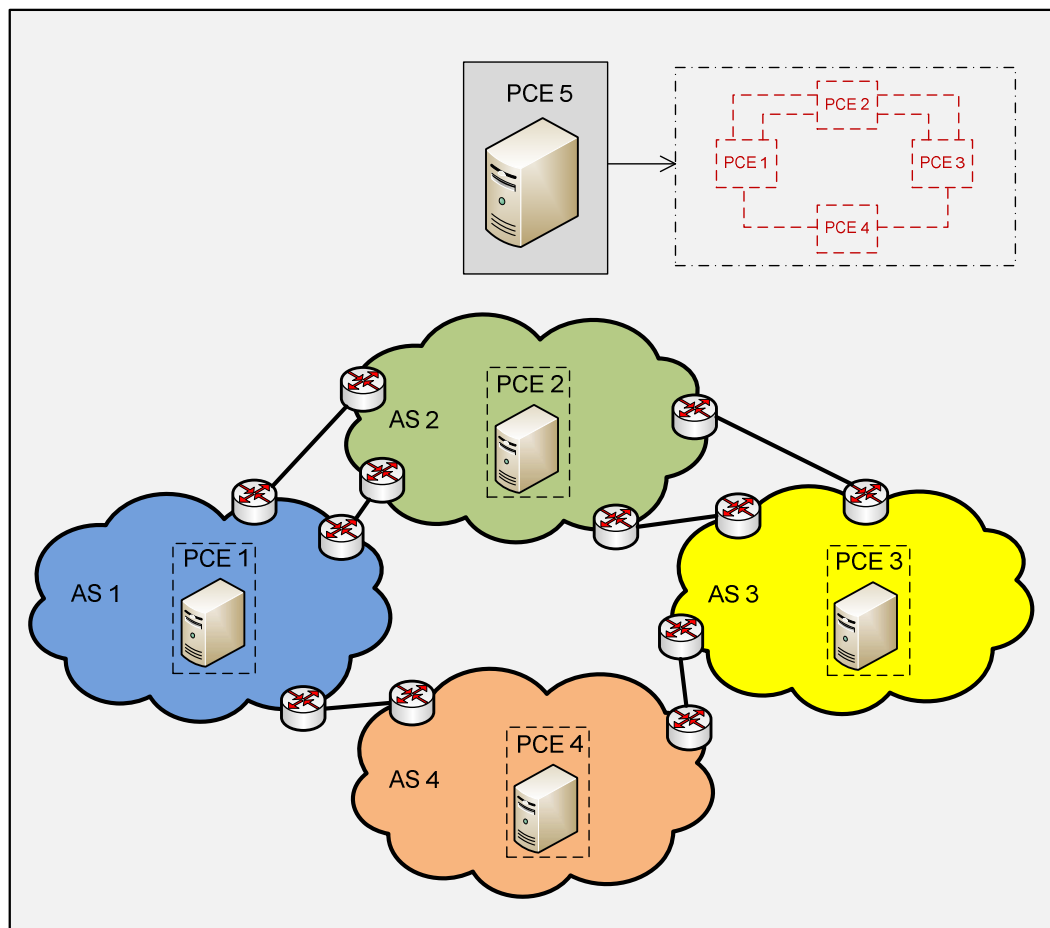


Figura 3.6 – Arquitectura PCE Jerárquica

En una arquitectura PCE jerárquica (Figura 3.6), un PCE padre mantiene un mapa de la topología multi-dominio en el que los nodos son los distintos dominios hijo y las conexiones son los enlaces inter-dominio. Dicho PCE padre contiene información de la identidad y la localización de los PCEs responsables de todos los dominios hijo (PCEs hijo), así como información referente a los recursos de ingeniería de tráfico de los enlaces inter-dominio.

Con el objetivo de preservar la confidencialidad de los diferentes dominios, el PCE padre no contiene ningún tipo de información acerca de las capacidades, recursos o conexiones dentro los dominios hijo. Además, cada PCE hijo sólo conoce la topología de su propio dominio y únicamente las identidades de sus dominios vecinos. Tampoco contienen información sobre la topología multi-dominio ni sus conexiones, exceptuando las que unen a su dominio con sus dominios vecinos.

Cada PCE hijo debe ser configurado con la dirección de su PCE padre. Es posible que un PCE hijo tenga más de un PCE padre. Así mismo, un PCE padre necesita conocer cada uno de los PCEs hijo de todos los dominios de su topología. Esta información puede ser configurada de antemano o bien aprendida cuando se establezca una conexión.

Los PCEs hijo puede enviar las identidades de sus dominios vecinos para que el PCE padre pueda obtener la información referente a las conexiones inter-dominio. Mediante IGP [25] y [26] pueden conocerse también las capacidades de ingeniería de tráfico de los enlaces inter-dominio. En teoría, un PCE padre podría participar en el protocolo IGP, pero esto violaría la confidencialidad y escalabilidad de los dominios hijo. Por ello, son los PCEs hijo quienes se encargan de reportar dicha información.

En la Figura 3.7 se muestra un ejemplo del proceso de cálculo en una arquitectura PCE jerárquica. El nodo A (origen) del dominio 1 solicita el cálculo de un camino hasta el nodo T (destino) del dominio 3 a su PCE local (PCE 1). Éste, al ver que el destino se encuentra fuera de su dominio, reenvía dicha solicitud al PCE padre (PCE 4). El PCE padre, que cuenta con un mapa global de conexiones de la topología, determina que existen cuatro caminos multi-dominio para alcanzar el destino. Estos caminos pueden estar sujetos a funciones objetivo [27] tales como coste mínimo, máximo ancho de banda, número mínimo de dominios a utilizar, etc. o bien a otro tipo de políticas como las relaciones entre operadores o el coste (en dinero) del camino [28]. A continuación pregunta a todos los PCEs hijo pertenecientes a los dominios involucrados en dichos caminos por las mejores rutas locales dentro de sus dominios: al PCE 1 le pregunta por rutas desde el origen (A) hasta los nodos de salida D y E, al PCE 3 por rutas desde los nodos de entrada N y O hasta el destino (T) y al PCE 2 (responsable de un dominio intermedio) por rutas desde sus nodos de entrada (F y G) hasta sus nodos de salida (L y M). Los PCEs hijo pueden reemplazar sus rutas locales con una path-key [29] de manera que mantengan la información de sus dominios oculta. Cuando recibe todas las respuestas, forma los caminos globales y determina cuál de ellos tiene un menor coste para proporcionárselo al PCE que lo solicitó (PCE 1). Éste, a su vez, se lo envía al PCC (nodo A) para su posterior establecimiento.

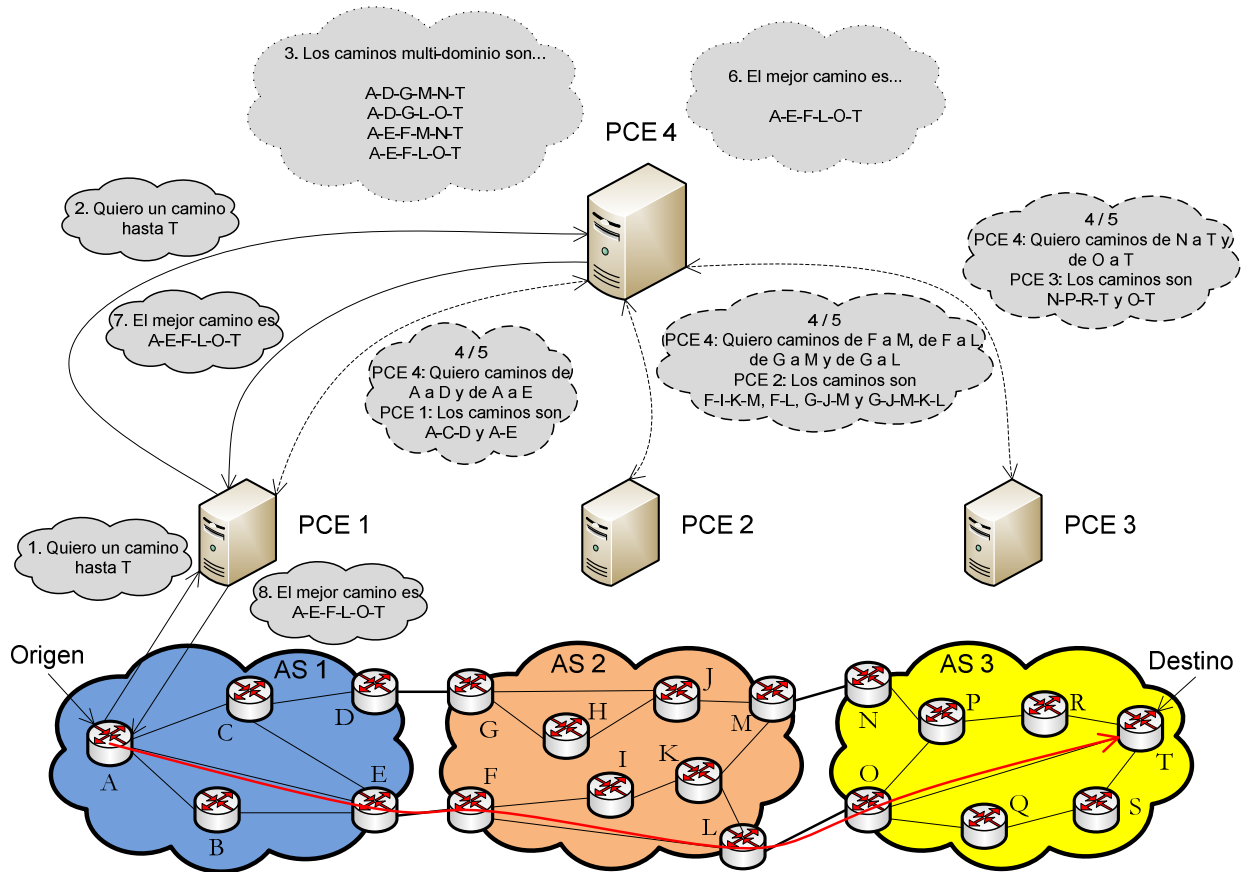


Figura 3.7 – PCE Jerárquico

Este procedimiento basado en una arquitectura jerárquica de los PCEs necesita una serie de extensiones del protocolo PCEP para funcionar. En primer lugar, debe ser posible que los PCEs hijo soliciten un camino al PCE padre indicando que sólo necesitan una secuencia de dominios en lugar de una camino completo. Esto permite a los PCEs hijo poder iniciar otros procedimientos como el per-domain o el BRPC. Por otro lado un PCE padre, antes de iniciar el cálculo de un camino, debería poder preguntar a un PCE hijo si el nodo destino se encuentra en su dominio. En el objeto *RP* del mensaje *Request* pueden indicarse estas especificaciones mediante el uso recomendado de dos banderas indicadas en [30].

Un PCE hijo debe ser capaz de indicar a un PCE padre si necesita o no que el PCE padre actúe como tal. Del mismo modo, un PCE padre podría indicar a un PCE hijo si puede actuar como padre cuando se le solicite. Se ha recomendado para este propósito la activación de dos banderas [30] en el objeto *Open* de los mensajes *Open* en la fase de inicialización de la sesión PCEP.

Por otra parte, como ya se ha indicado antes, los PCEs hijo deben informar al PCE padre de su dirección y los identificadores de sus dominios vecinos, así como de la información referente a los enlaces inter-dominio (identificadores del enlace y de los nodos de borde origen y destino, información de ingeniería de tráfico, etc.). Esta información podría ser enviada en mensajes *Notification*. Estas y otras extensiones pueden encontrarse en [30].

4. Comparativa de las soluciones basadas en PCE para entornos multi-dominio

En este proyecto se ha trabajado sobre una implementación del PCE para hacerla funcionar en entornos multi-dominio. Concretamente, se han implementado tres de las cuatro soluciones explicadas en el capítulo anterior: *Per-domain Path Computation*, *Backward-Recursive PCE-based Computation* y *PCE Jerárquico*. Para validar su funcionamiento, estudiar sus prestaciones y comparar las tres soluciones implementadas en términos de probabilidad de bloqueo, se han realizado una serie de experimentos.

El capítulo está formado por tres secciones. La sección 4.1 presenta una descripción detallada del escenario de trabajo, en la sección 4.2 se analizan los tres protocolos en función de sus resultados de probabilidad de bloqueo y, por último, en la sección 4.3 se proponen y se comparan las soluciones implementadas a los problemas presentados en la sección 4.2.

4.1. Descripción del escenario

Con el fin de comparar las prestaciones de las tres soluciones se ha utilizado una topología de red multi-dominio real (Figura 4.1). Dicha topología es la *Red Troncal China* [31], que cuenta con cinco ASes con diferente número de nodos. En la Tabla 4.1 se muestran sus características. Además, se han evaluado los protocolos en otra topología, recogiendo en el anexo A toda la información.

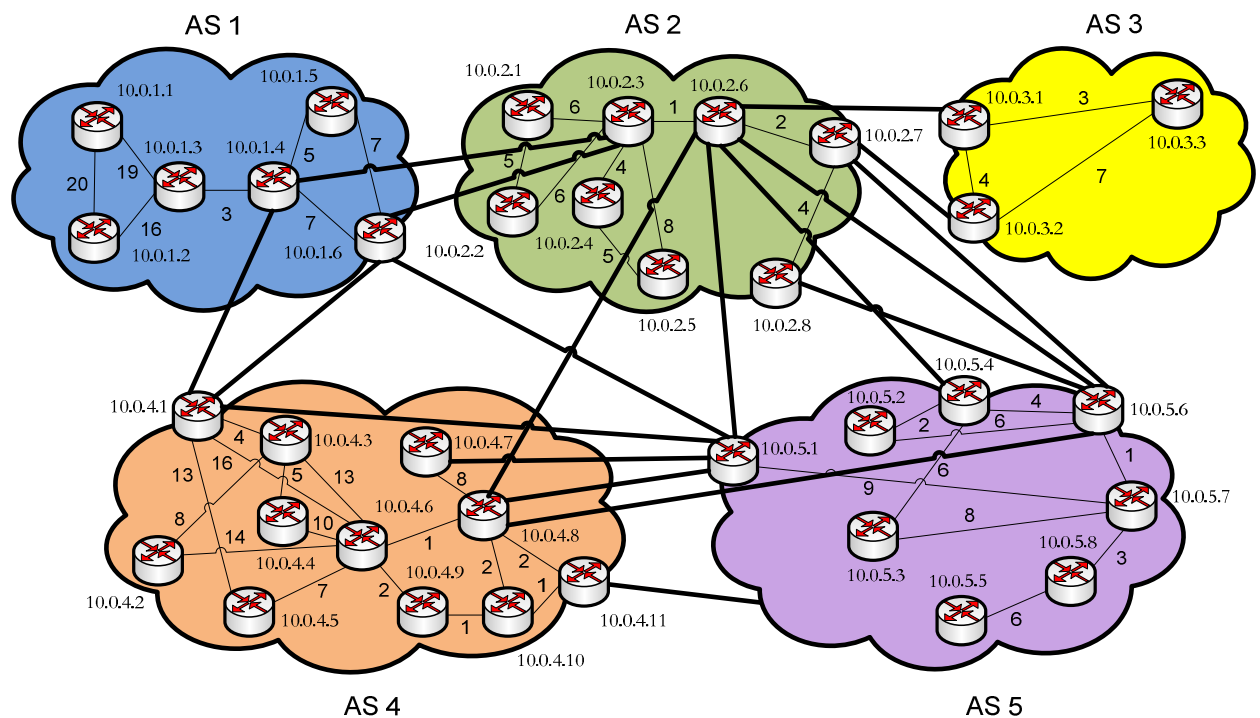


Figura 4.1 - Red Troncal China [31]

Tabla 4.1 – Características de la *Red Troncal China* (36 nodos)

Número de AS	Número de nodos	Número de enlaces (bidireccionales)	Grado medio de los nodos
1	6	7	2.33
2	8	10	2.5
3	3	3	2
4	11	16	2.91
5	8	9	2.25

Los PCEs están formados por dos módulos diferenciados (Figura 4.2). Uno de ellos (PCEP) se encarga de la comunicación con PCCs y PCEs y el otro es el encargado del cálculo de caminos. Este último, a su vez, está formado por otros tres módulos, los cuales realizan las funciones de enrutamiento, validación de restricciones y elección de longitud de onda. Así, un PCE opera de la siguiente forma: el módulo de validación de restricciones evalúa un camino calculado previamente por el módulo de enrutamiento y, a continuación, el módulo de elección de longitud de onda elige una disponible. De esta manera, el módulo de validación de restricciones puede rechazar caminos antes de la elección de longitud de onda hasta que se encuentre un camino adecuado o se supere el número máximo de intentos.

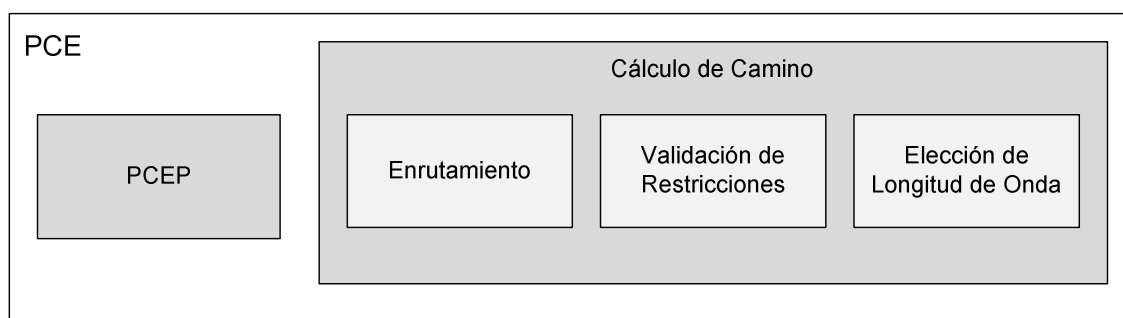


Figura 4.2 - Composición de un PCE

En el módulo de enrutamiento se ha usado el algoritmo Dijkstra K-shortest path con $K = 5$, en donde se calcula el mejor camino posible entre el origen y el destino y no se permite que los nodos intermedios se repitan, es decir, que haya bucles. En los experimentos realizados, el coste utilizado por el algoritmo ha sido el número de tramos de fibra de los enlaces (spans).

Para el módulo de elección de longitud de onda se ha usado el algoritmo First-Fit (FF), que selecciona la primera longitud de onda disponible (empezando desde la 0) hasta agotar los recursos. Se asume que el PCE es la única entidad capaz de realizar la elección de longitudes de onda para los caminos en la red y

que contiene información de los recursos asignados y disponibles en todos los enlaces en la TED. Para almacenar esta información se ha utilizado un servidor MySQL.

En cuanto a las restricciones físicas de los enlaces, se han tenido en cuenta la relación señal a ruido ($OSNR_{min}$) y la dispersión del modo de polarización (PMD). En menor medida, se han considerado otros efectos mediante un margen $OSNR_{res}$ en dB. El estándar ITU-T G.680 proporciona una ecuación para el cálculo de la OSNR en el receptor de un camino óptico [32]. Una versión simplificada de esta ecuación es la siguiente:

$$OSNR = P_{out} - \alpha L - 10 \log_{10} N + 58 - F, \quad (1)$$

en donde OSNR es la relación señal a ruido en dB, P_{out} es la potencia de señal en dBm a la salida del último amplificador, α es el factor de atenuación en dB/Km, L es la longitud de un tramo de fibra o span, N es el número total de spans y F es el factor de ruido de los amplificadores ópticos en dB. El valor de 58 en (1) es igual a $10 \log_{10}(hf_i \Delta f_0)$, siendo h la constante de Planck, f_i la frecuencia óptica en Hz y Δf_0 el ancho de banda de referencia en Hz. Se considera este valor válido para todos los canales al ser la diferencia para todas las longitudes de onda insignificante. De esta manera, el PCE evalúa la siguiente condición para todos los caminos después del módulo de enrutamiento

$$OSNR > OSNR_{min} + OSNR_{res}, \quad (2)$$

en donde $OSNR_{min}$ es la relación señal a ruido requerida y $OSNR_{res}$ el margen acumulado de otras restricciones físicas. A partir de (1), el módulo de validación de restricciones puede obtener el número máximo de spans N_{max} que satisface (2).

Los autores de [33] muestran la siguiente ecuación simplificada para la PMD

$$PMD = B \sqrt{\sum_{K=1}^N D_{PMD}^2} L, \quad (3)$$

siendo B la tasa de símbolos del canal y D_{PMD} el valor de PMD para cada tramo de fibra o span. De forma similar al límite impuesto por la OSNR, se puede determinar un valor de N_{max} para el límite impuesto por la PMD en (3) y elegir el valor mínimo de los dos.

En cuanto a los amplificadores ópticos, todos ellos cuentan con una ganancia G que compensa las pérdidas de transmisión de un tramo de fibra ($G = \alpha L$). También se aplica control de ganancia y todos tienen un factor de ruido similar.

Por tanto, los caminos ópticos dependen únicamente en (1) y (3) del parámetro N , lo que simplifica el proceso a comprobar los límites de OSNR y PMD, y determinar el valor de N para los distintos caminos para compararlo con N_{max} .

El número máximo de longitudes de onda con que cuentan los enlaces ópticos es de $M = 80$. Los valores para las restricciones físicas son los siguientes: $P_{\text{out}} = 4$ dBm, $\alpha = 0.35$ dB/Km, $L = 80$ Km, $\text{OSNR}_{\text{res}} = 1.5$ dB, $\text{OSNR}_{\text{min}} = 12$ dB, $B = 10$ Gbps y $D_{\text{PMD}} = 0.2 \cdot 10^{-12}$ s/Km^{0.5}.

Para los experimentos se ha considerado que en cada dominio todas las solicitudes de caminos se realizan desde un mismo módulo denominado PCC Fuente. De esta manera (véase la Figura 4.3), se tendría un dominio origen con 6 nodos numerados del 1 al 6. Suponiendo que existen 4 dominios en toda la red (incluyendo el dominio origen), y que todos ellos cuentan con 6 nodos, se pueden generar valores aleatorios uniformemente para cada solicitud del nodo origen (entre 1 y 6), del dominio destino (entre 1 y 4) y del nodo destino (entre 1 y 6).

La sesión establecida entre el PCC Fuente y el PCE es permanente, es decir, todas las solicitudes se envían en la misma sesión PCEP, que se mantiene activa mediante mensajes *Keepalive*.

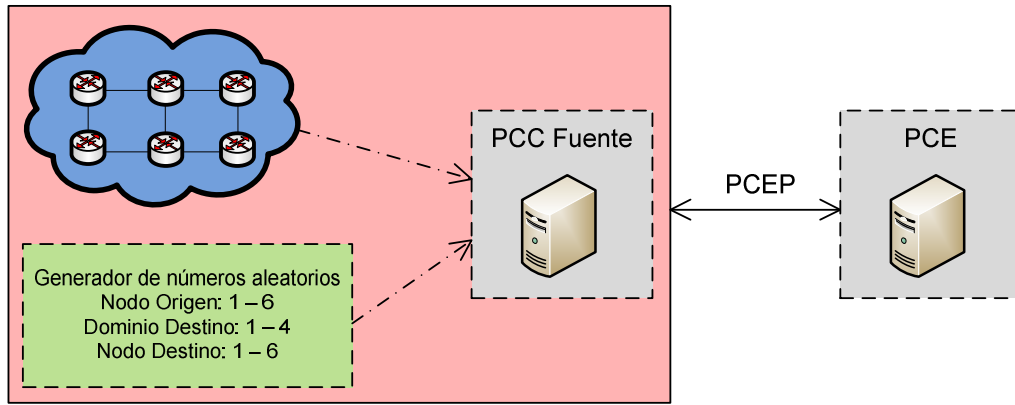


Figura 4.3 – Generación de tráfico en un PCC Fuente

La llegada de solicitudes sigue un proceso de Poisson con tasa de llegadas exponencial en cada PCC Fuente. El tiempo de servicio de cada solicitud se distribuye exponencialmente. De esta manera, se tiene

$$A = aN = \frac{\lambda}{\mu}$$

siendo A el tráfico total ofrecido en Erlangs, a el tráfico ofrecido por cada nodo en Erlangs, N el número de nodos del dominio, λ la tasa media de llegadas y μ la tasa media de servicio.

Los PCEs y PCCs se encuentran en el mismo servidor, por lo que los retardos que se han tenido en cuenta son los correspondientes al protocolo PCEP y a los algoritmos de enrutamiento y elección de longitud de onda. En una red real existen otros retardos, como el introducido por el plano de control o el retardo de propagación debido a que los PCEs se encuentran en sitios diferentes, que no han sido incluidos en estos experimentos.

En el diagrama de la Figura 4.4 se puede observar la implementación realizada para el per-domain. Al realizarse dicho procedimiento durante el proceso de señalización (véase la sección 3.2), la reserva de longitud de onda y establecimiento del camino se han realizado en los PCEs de cada dominio de la cadena por motivos de implementación.

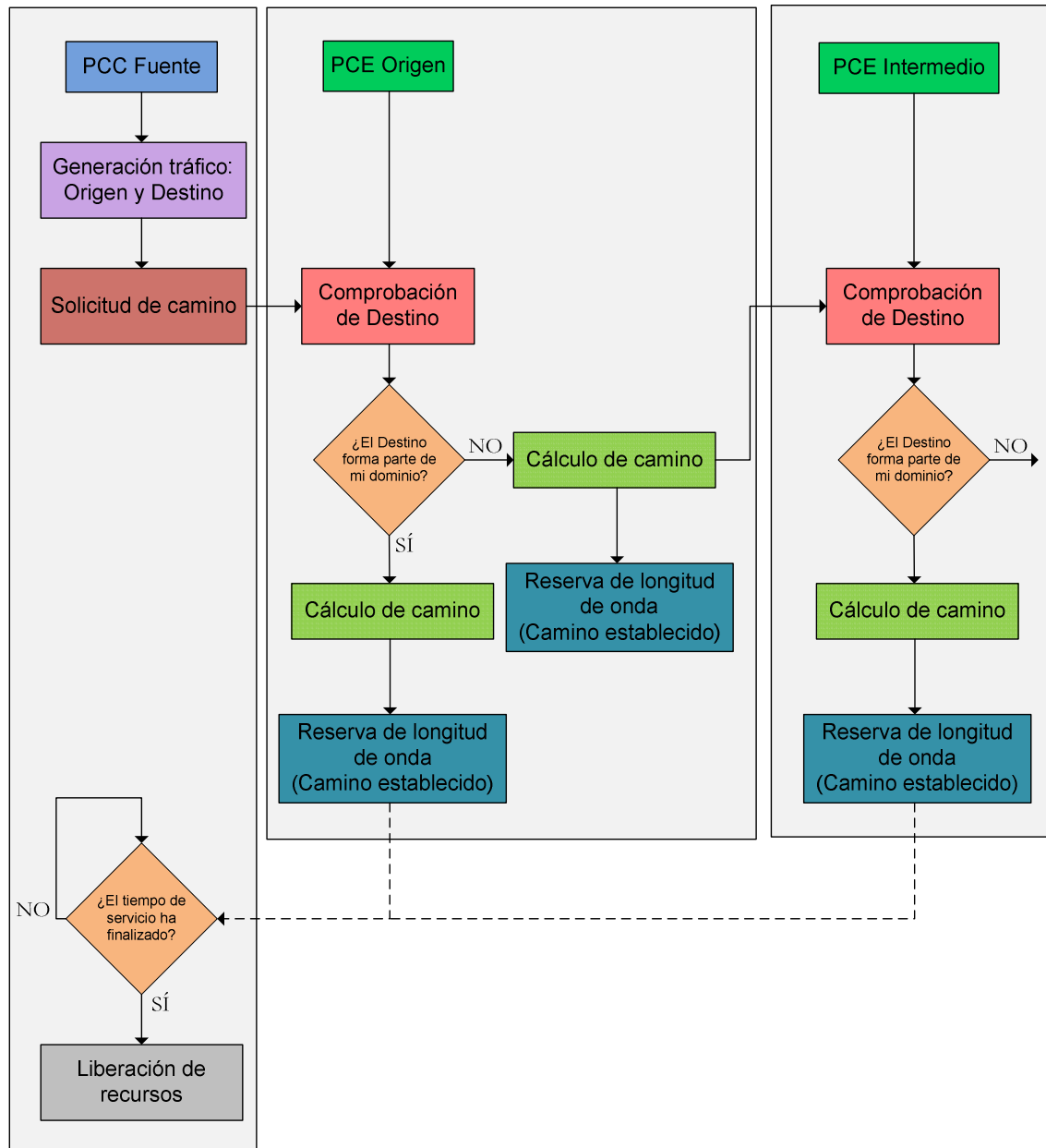


Figura 4.4 - Implementación del Per-domain Path Computation

En la Figura 4.5 puede verse el diagrama que representa el proceso seguido por el BRPC donde, tras alcanzar la solicitud el PCE del dominio destino, los distintos PCEs de la cadena colaboran entre sí en la consecución del camino (véase la sección 3.3). El módulo *Cálculo de camino* realiza las funciones de enrutamiento, validación de restricciones y elección de longitud de onda. Cuando el PCC Fuente recibe el camino calculado debe proceder a la reserva de longitudes de onda en los distintos dominios que forman parte del camino. Posteriormente se activa un temporizador para simular el tiempo de servicio del camino y, cuando el temporizador expira, se liberan las longitudes de onda previamente reservadas.

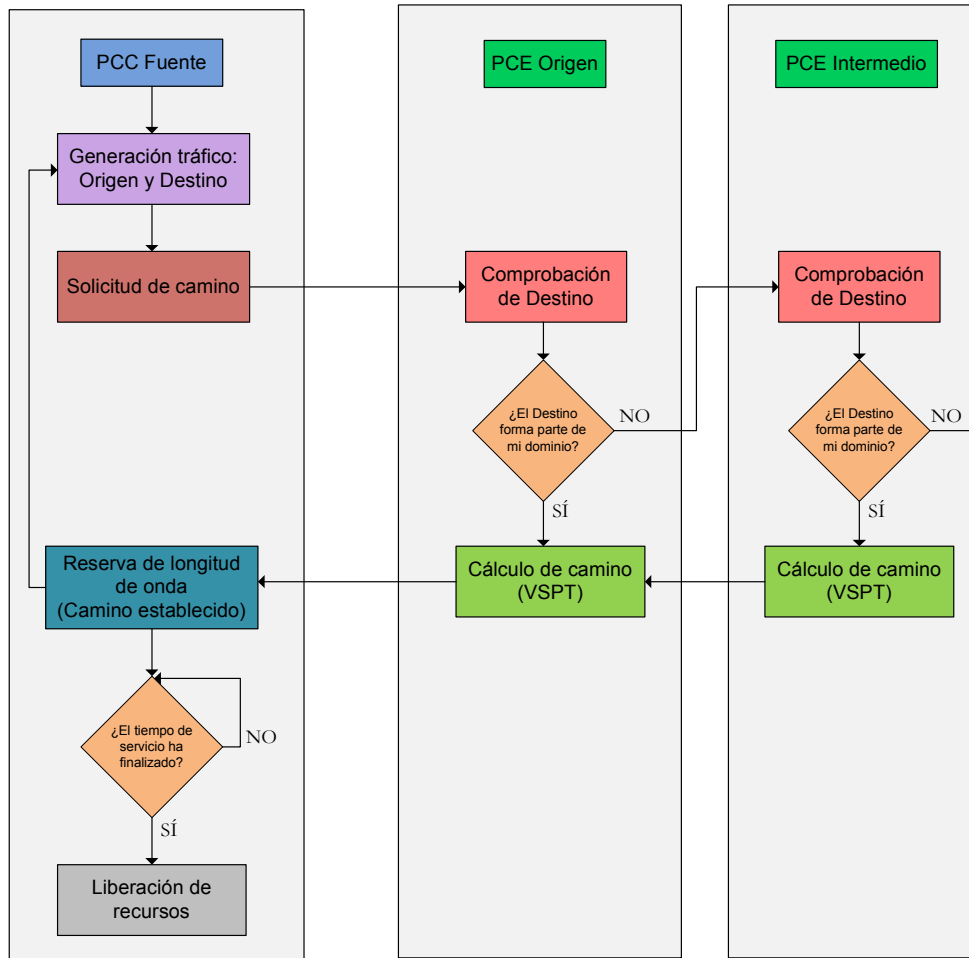


Figura 4.5 - Implementación del BRPC

Para el transporte de las longitudes de onda en los mensajes *Reply* se ha utilizado subobjetos *Label ERO Subobjects*, cuyas especificaciones se encuentran en [7] y [33], a continuación de cada subobjeto *ERO* (que representa un nodo).

En el caso de que no haya ninguna longitud de onda disponible o que no se pueda asignar debido a las restricciones físicas de los enlaces, se ha utilizado dentro del objeto *NO-PATH* un subobjeto *TLV* denominado *NO-PATH-VECTOR TLV* [15] con el fin de proporcionar información adicional sobre el motivo que ha causado una respuesta negativa. Los bits asignados en el campo de banderas en este subobjeto han sido el 21 si no se ha podido encontrar un camino, el 22 si existen restricciones físicas y el 23 si no se ha podido encontrar ninguna longitud de onda disponible.

En la Figura 4.6 puede observarse el diagrama correspondiente al procedimiento H-PCE. Como se explicó en la sección 3.4, el PCC Fuente envía una solicitud de camino al PCE de su dominio y éste comprueba si el camino solicitado es intra-dominio o inter-dominio. En este último caso, reenvía la solicitud al PCE padre, el cual se encarga de calcular el camino.

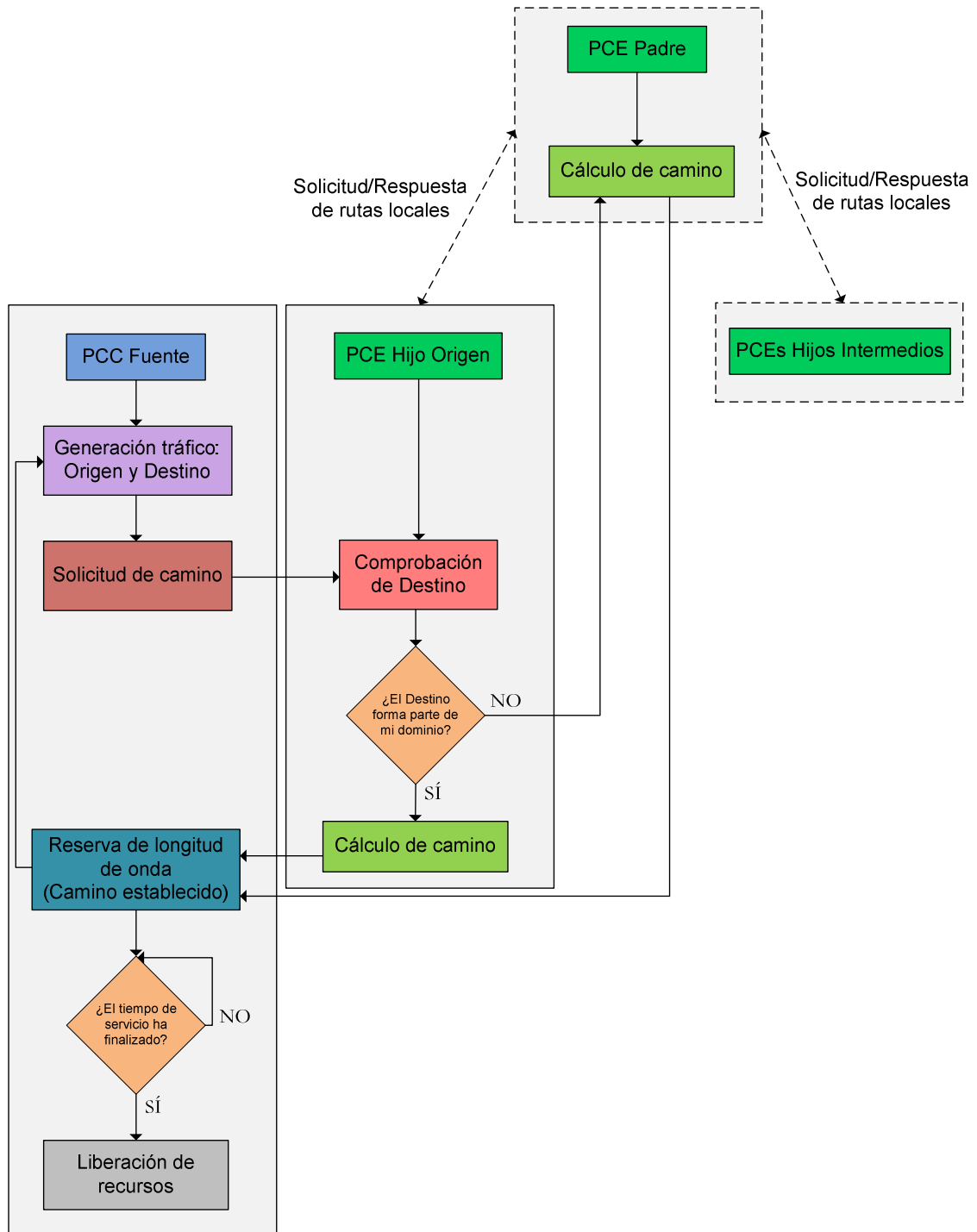


Figura 4.6 - Implementación del H-PCE

4.2. Probabilidad de bloqueo

En los experimentos se han hecho un total de 20000 peticiones por cada uno con el objetivo de conseguir resultados estables. De esta manera, en la Figura 4.7 se puede ver cómo se estabiliza el bloqueo para el protocolo per-domain.

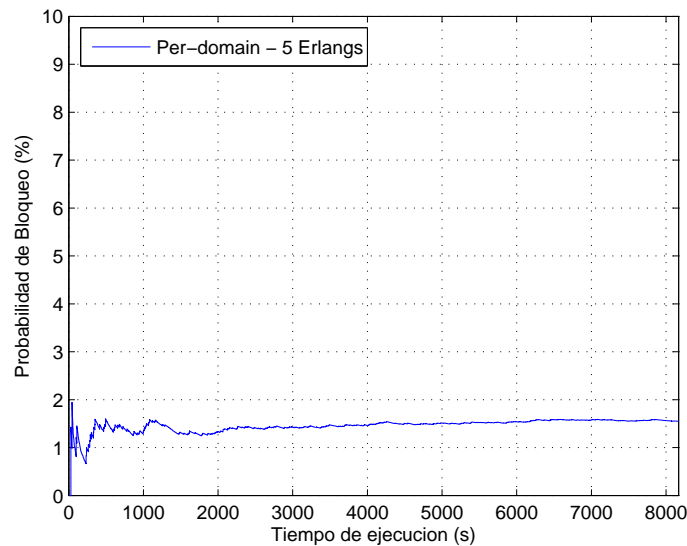


Figura 4.7 - Estabilidad para per-domain

Esto también se ha comprobado para el BRPC (Figura 4.8) y el H-PCE (Figura 4.9) para distintos valores de K y un tráfico ofrecido por nodo diferente.

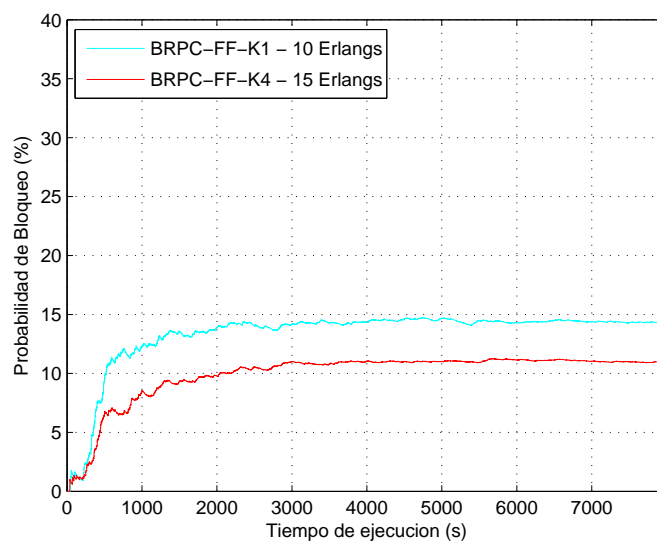


Figura 4.8 - Estabilidad para BRPC

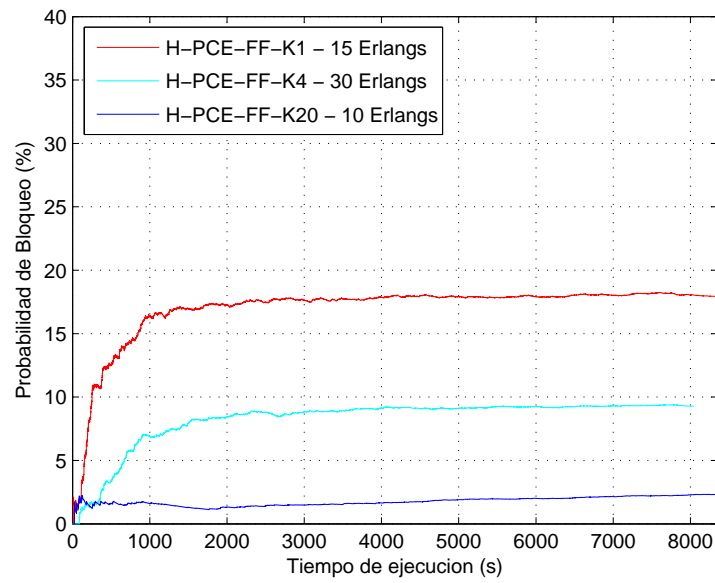
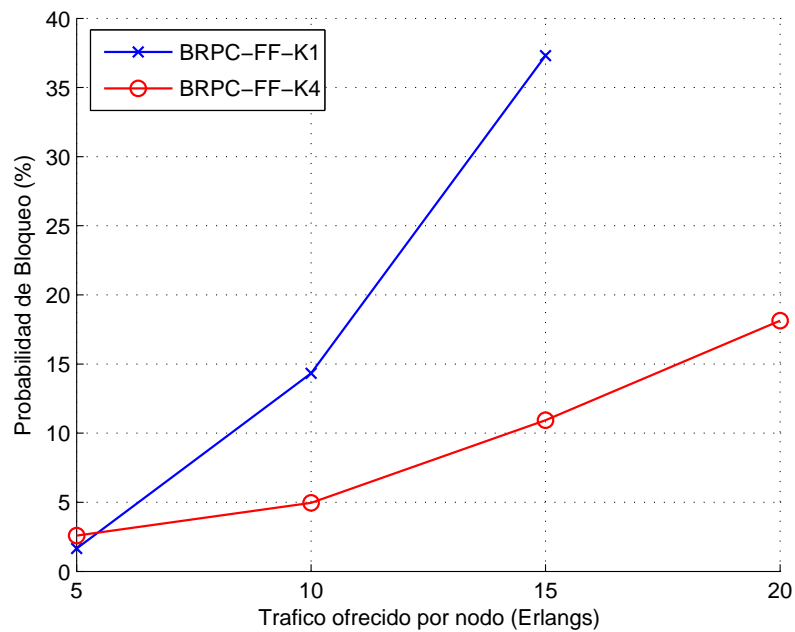


Figura 4.9 - Estabilidad para H-PCE

El número máximo de secuencias de dominios que se pueden comprobar en la topología de la Figura 4.1 para el cálculo de un camino mediante el procedimiento BRPC es de $K_{max} = 4$. Se ha querido comprobar que, cuanto mayor sea K , menor probabilidad de bloqueo se debería obtener. Por tanto, utilizando el algoritmo First-Fit (FF) en la elección de longitud de onda, se han conseguido los siguientes resultados (Figura 4.10) en términos de probabilidad de bloqueo para $K = 1$ y $K = K_{max} = 4$.

Figura 4.10 - Probabilidad de bloqueo para BRPC variando K con el algoritmo FF

Se observa que para un tráfico ofrecido de 5 Erlangs por nodo la probabilidad de bloqueo es mayor para $K = 4$ que para $K = 1$. Esto es así debido a un tipo de bloqueo adicional (Figura 4.11) que aparece cuando el PCC Fuente intenta reservar una longitud de onda pero esta ya ha sido reservada por otro PCC Fuente. En el módulo de elección de longitudes de onda del PCE el recurso estaba disponible, no así en la reserva del camino por parte del PCC Fuente. Esto está condicionado por el tiempo que la solicitud ha tardado en ser atendida y devuelta al PCC Fuente. Al tener que comprobar $K = 4$ caminos, el tiempo de cómputo en los PCEs crece (Figura 4.11) dando como resultado este tipo de bloqueo que se denominará “robo de longitud de onda”.

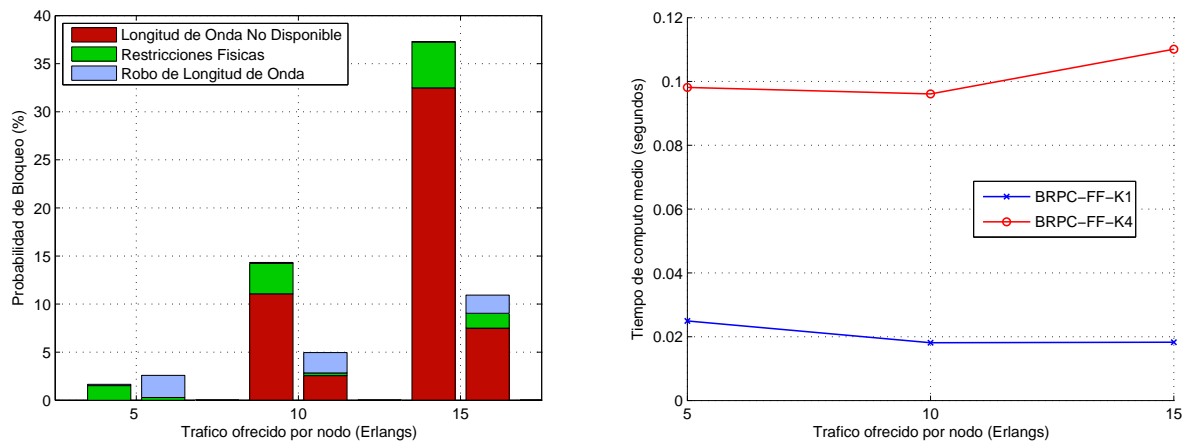


Figura 4.11 - Tipos de bloqueo (izda.) y tiempos de cómputo (dcha.) para BRPC variando K con el algoritmo FF

Con respecto al H-PCE, el número máximo de caminos que el PCE padre puede comprobar asciende hasta $K_{max} = 20$. Es conveniente indicar que en el BRPC, K hace referencia a caminos multi-dominio (secuencias de dominios) y en el H-PCE a todos los caminos posibles en función de los nodos de borde y las conexiones inter-dominio (véase la sección 3.4). De este modo, en ambos casos se comprueban todas las posibilidades y podría hacerse una comparativa entre los dos protocolos de forma justa.

Los resultados para el H-PCE (Figura 4.12) son similares a los del BRPC. Cuanto mayor es K , mejores resultados se obtienen o se deberían obtener si se eliminara el bloqueo por “robo de longitud de onda” explicado anteriormente.

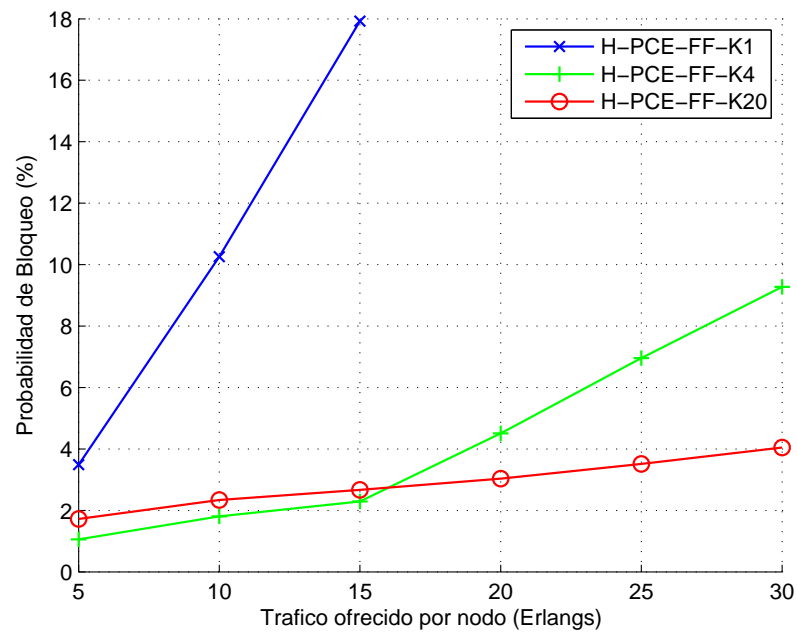


Figura 4.12 - Probabilidad de bloqueo para H-PCE variando K con el algoritmo FF

Se ve claramente que para 5, 10 y 15 Erlangs los resultados no son los esperados, lo cual se comprueba observando la Figura 4.13, en donde el tiempo de cómputo para $K = 20$ es mucho mayor que en los otros dos casos, lo que provoca un mayor bloqueo por “robo de longitud de onda”.

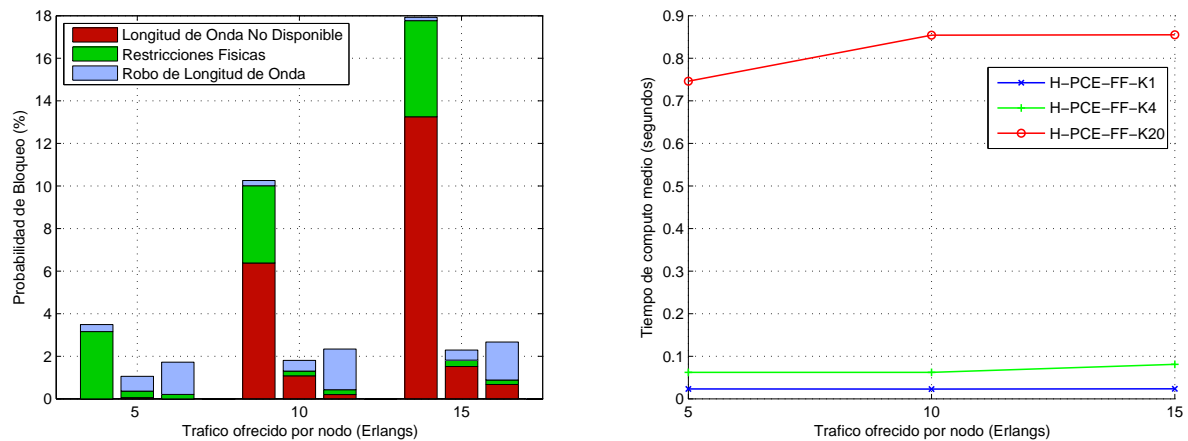


Figura 4.13 - Tipos de bloqueo (izda.) y tiempos de cómputo (dcha.) para H-PCE variando K con el algoritmo FF

Para el Per-domain Path Computation se han obtenido los resultados de la Figura 4.14. Para este caso, $K_{max} = 1$.

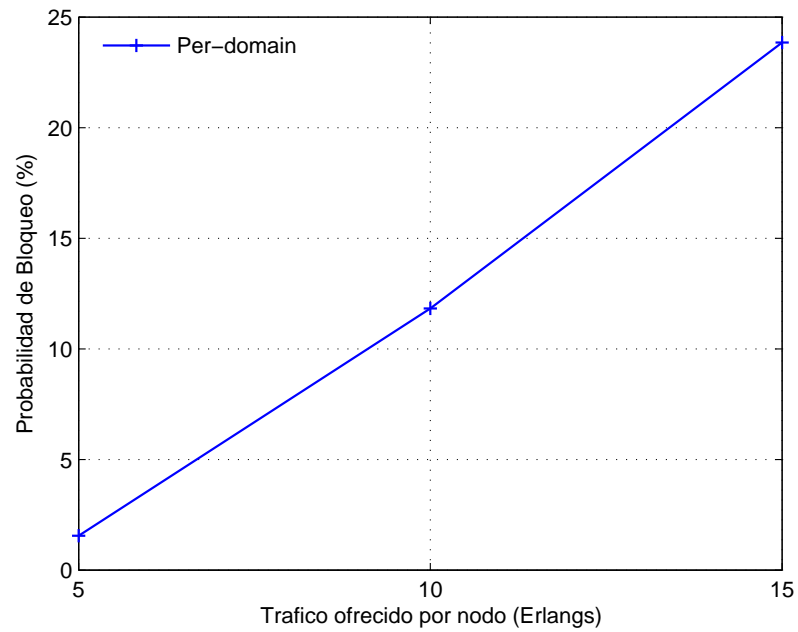


Figura 4.14 - Probabilidad de bloqueo para per-domain con el algoritmo FF

Con el objetivo de eliminar el bloqueo por “robo de longitud de onda” en la reserva de recursos, se ha optado por implementar dos soluciones diferentes, las cuales se explican en la sección siguiente.

4.3. Soluciones propuestas para la mejora de la probabilidad de bloqueo por “robo de longitud de onda”

Para solventar el problema del bloqueo por “robo de longitud de onda” primero se optó por una solución sencilla consistente en modificar el algoritmo First-Fit de asignación de recursos. Se ha llamado a este nuevo algoritmo Round-Robin First-Fit (RR-FF). De manera local (en cada PCE) se guarda el último valor de longitud de onda elegido (hasta $M = 80$) y en la siguiente solicitud se comienza a buscar el valor disponible a partir de éste. Como se puede apreciar para el BRPC en la Figura 4.15, se consigue reducir el bloqueo por “robo de longitud de onda” pero no hasta hacerlo desaparecer, como se desea.

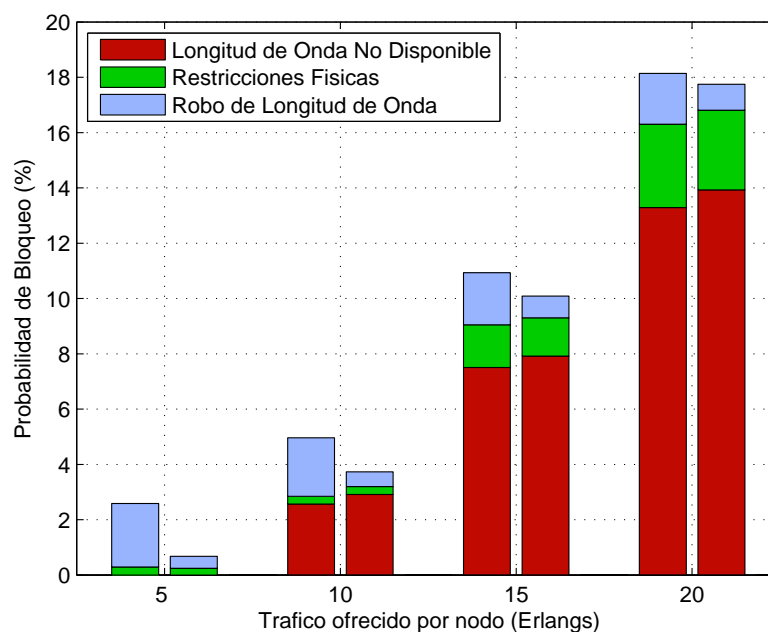


Figura 4.15 - Tipos de bloqueo para BRPC con los algoritmos FF (izda.) y RR-FF (dcha.)

En el caso del H-PCE el resultado es similar (Figura 4.16). La mejora es apreciable, pero lo ideal sería contar con una solución que hiciera desaparecer este bloqueo por completo. Obsérvese que en ambos casos (BRPC y H-PCE), al reducirse el bloqueo por “robo de longitud de onda”, se reservan más recursos y, por tanto, hay menos disponibles para su elección en el módulo correspondiente del PCE. Esto produce un ligero aumento del bloqueo en la elección de longitudes de onda.

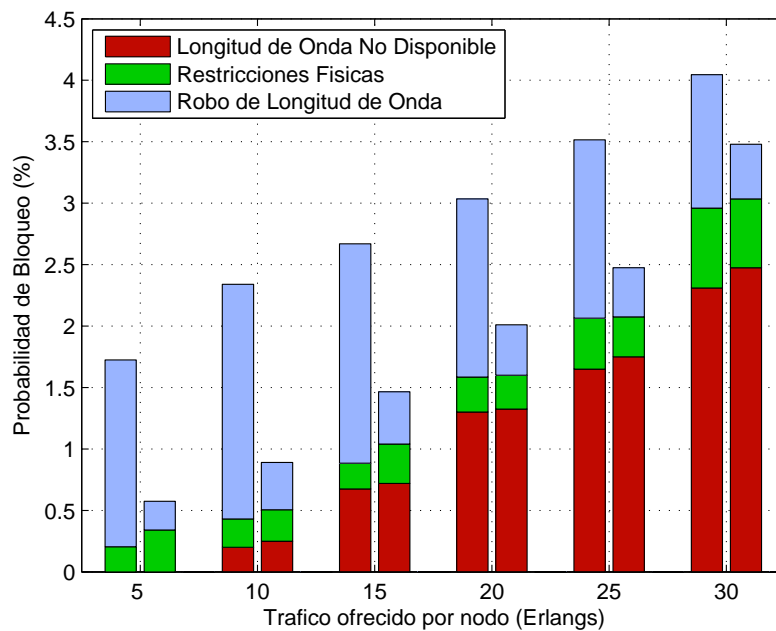


Figura 4.16 - Tipos de bloqueo para H-PCE con los algoritmos FF (izda.) y RR-FF (dcha.)

La segunda solución se puede encontrar en [35], en donde sus autores proponen un mecanismo de pre-reserva de los recursos con el objetivo de evitar el bloqueo por “robo de longitud de onda”. Mediante dicho mecanismo, cuando existe una solicitud de camino, el PCE pre-reserva los recursos en su TED por un tiempo especificado por el PCC (T_{res}). Una vez que dicho tiempo expira los recursos son liberados, de tal manera que puedan asignarse a otras solicitudes de caminos. Para esta implementación se han incluido dentro del mensaje *Request* un objeto *PCC-ID-REQ* (cuyas características se presentan en [36]) para especificar la dirección IP del PCC que solicita la reserva temporal de recursos, así como un objeto *RESERVATION* (especificado en [35]) indicando el tiempo (T_{res}) que los recursos deberán estar pre-reservados.

Se ha comprobado que cuando el valor del temporizador es suficientemente alto ($T_{res} = 3s$ para los experimentos de este proyecto), el bloqueo por “robo de longitud de onda” desaparece por completo. Si se hubiera utilizado un valor menor, la probabilidad de que siga existiendo dicho bloqueo crecería, como puede intuirse, al estar los recursos reservados un menor tiempo. En la Figura 4.17 se puede ver la mejora introducida con el mecanismo de pre-reserva (PR) para el BRPC con respecto a los algoritmos FF y RR-FF.

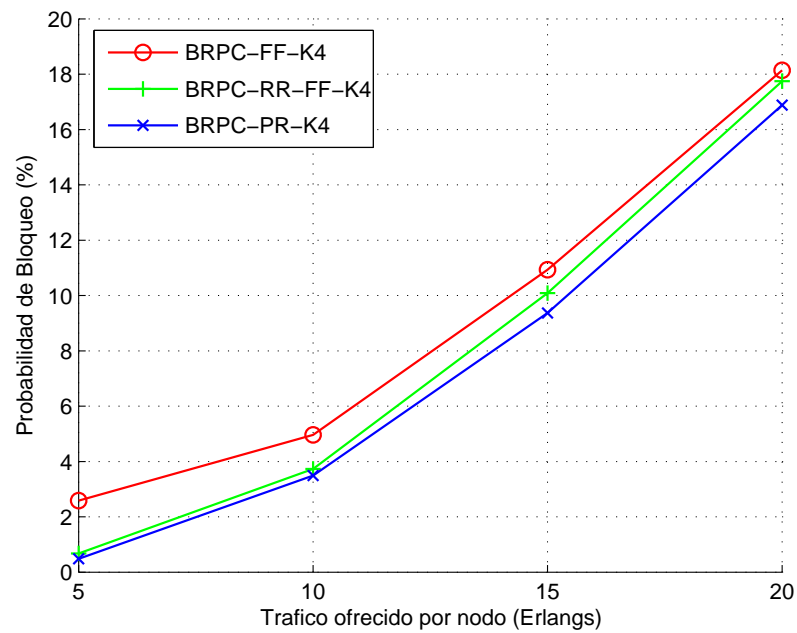


Figura 4.17 - Probabilidad de bloqueo para BRPC con los algoritmos FF y RR-FF y PR

Se aprecia una reducción de la probabilidad de bloqueo usando el mecanismo PR. El bloqueo por “robo de longitud de onda” se elimina completamente, como se puede observar en la Figura 4.18.

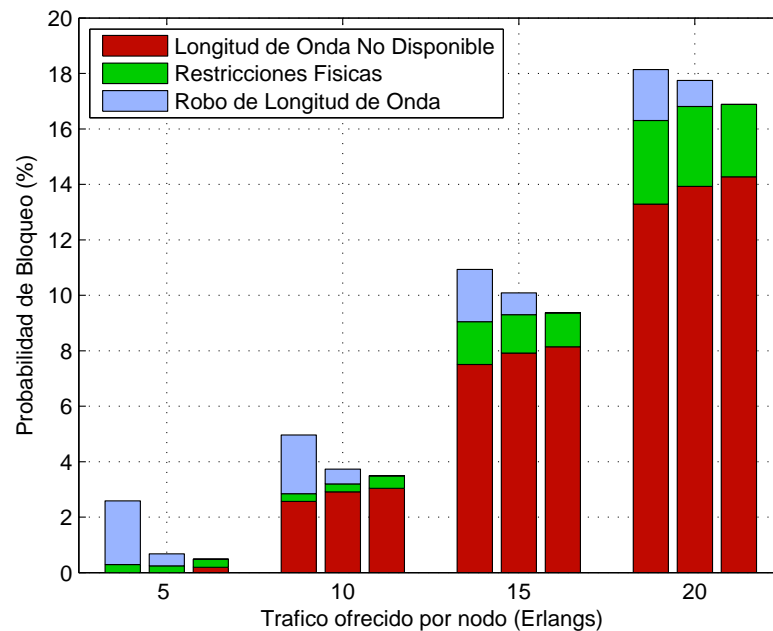


Figura 4.18 - Tipos de bloqueo para BRPC con los algoritmos FF (izda.) y RR-FF (centro) y PR (dcha.)

Para el H-PCE se pueden observar los mismos resultados. La mejora introducida se ve claramente en la Figura 4.19.

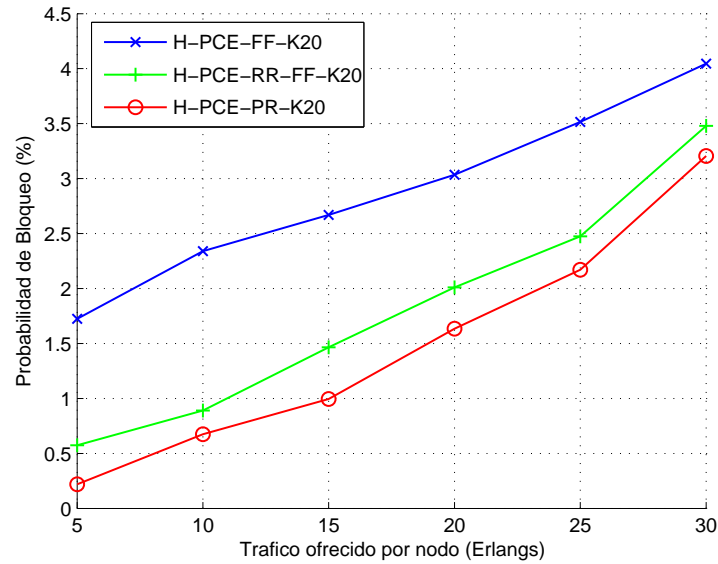


Figura 4.19 - Probabilidad de bloqueo para H-PCE con los algoritmos FF y RR-FF y PR

El bloqueo por “robo de longitud de onda” es eliminado por completo, apreciándose esta importante reducción en la Figura 4.20. De nuevo se observa para ambos protocolos un ligero aumento del bloqueo en la elección de longitudes de onda por estar todas asignadas.

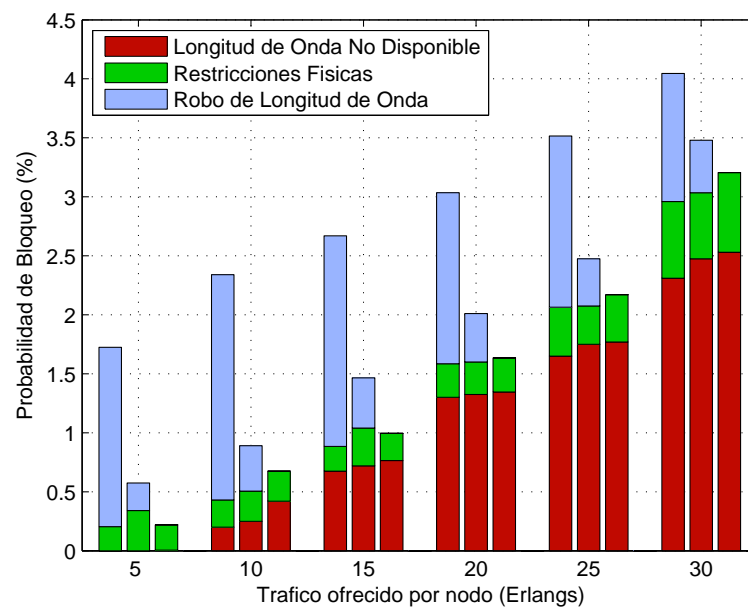


Figura 4.20 - Tipos de bloqueo para H-PCE con los algoritmos FF (izda.), RR-FF (centro) y PR (dcha.)

En la Figura 4.21 se presenta una comparativa en términos de probabilidad de bloqueo para los tres protocolos utilizando el mecanismo de pre-reserva y para distintos valores de K .

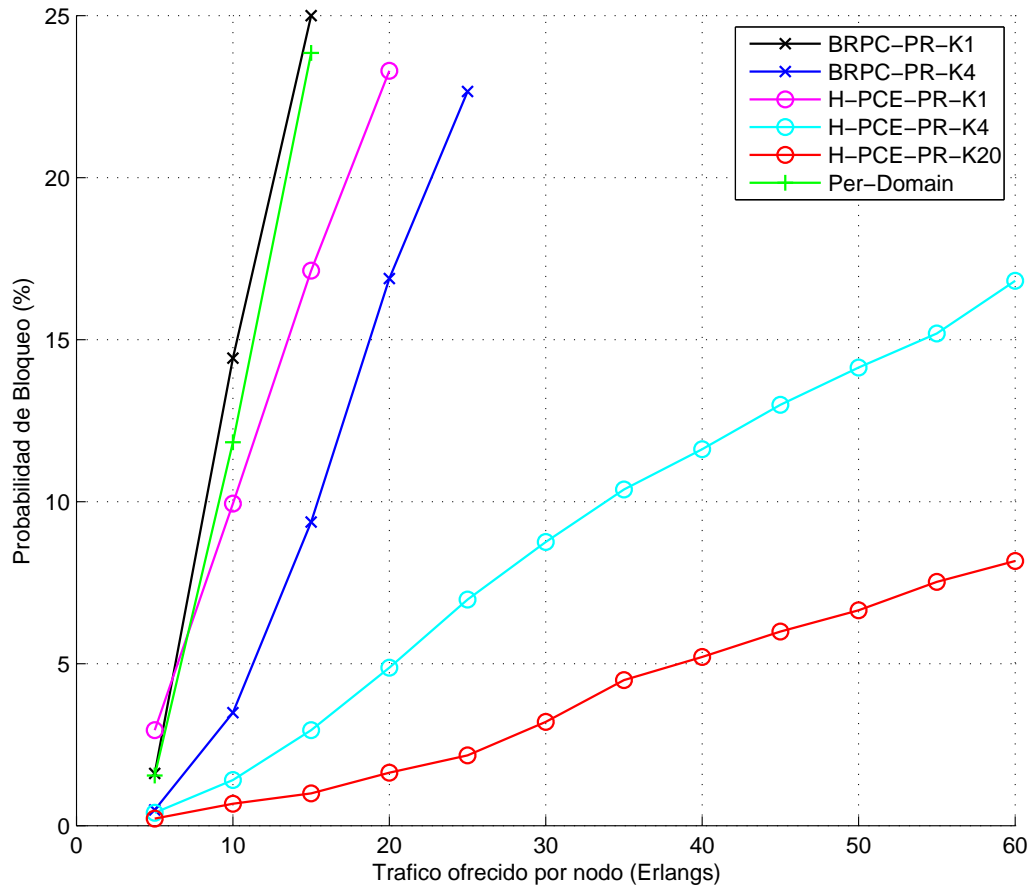


Figura 4.21 - Comparativa en términos de probabilidad de bloqueo con el mecanismo PR

El H-PCE presenta los mejores resultados con respecto a los otros dos protocolos. Para $K = K_{max}$ y un tráfico ofrecido de 20 *Erlangs* por nodo, la probabilidad de bloqueo en el H-PCE ($K = 20$) no llega al 2.5%, mientras que para el BRPC ($K = 4$) es del 17%.

El tiempo de cómputo, como se puede observar en la Figura 4.22, es considerablemente mayor en el caso del H-PCE con $K = 20$. Para $K = 4$ los resultados, aunque peores en términos de probabilidad de bloqueo respecto a $K = 20$, siguen siendo indudablemente mejores que los del BRPC y el per-domain y, además, su tiempo de cómputo es menor.

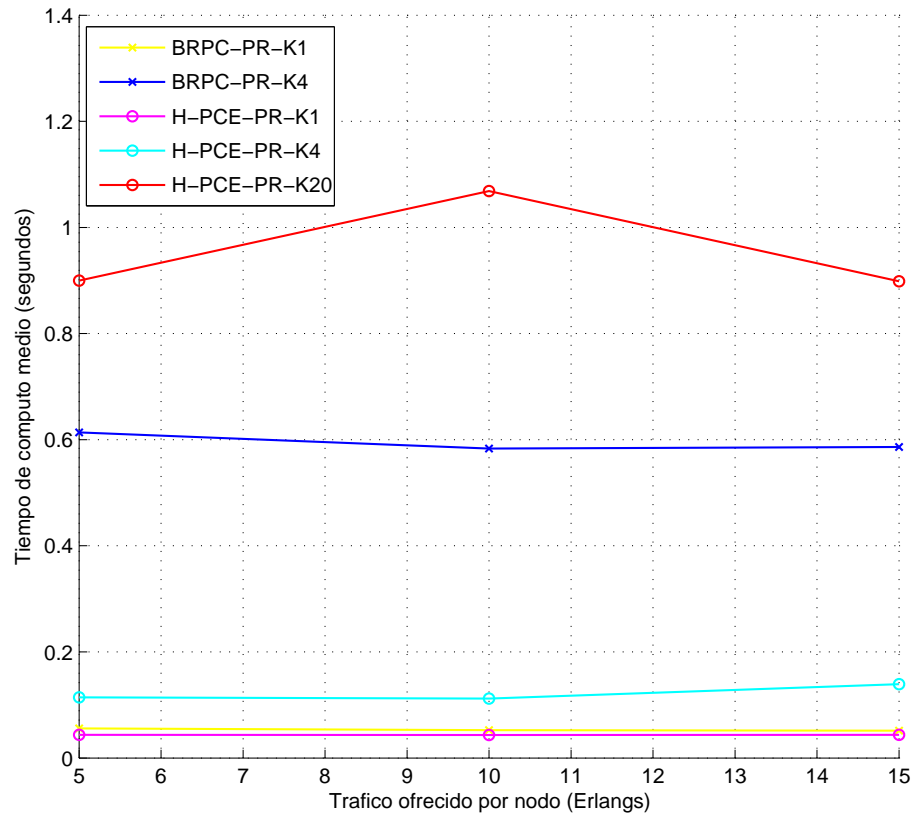


Figura 4.22 - Comparativa de los tiempos de cómputo con el mecanismo PR

Es importante destacar también que para el BRPC y $K = 1$ la probabilidad de bloqueo es ligeramente peor que para el per-domain (Figura 4.21). Esto es así, como ya se explicó en la sección 3.3, porque la topología con la que se han realizado los experimentos cuenta con múltiples conexiones entre los distintos dominios y el proceso de cálculo se complica. Además, para la consecución de las secuencias de dominios se ha utilizado también el algoritmo *K-shortest path*, por lo que para $K = 1$ y los mismos dominios de origen y destino, la secuencia siempre es la misma y el bloqueo, por tanto, aumenta.

5. Conclusiones y futuras líneas de trabajo

Este capítulo está dividido en dos secciones. La sección 5.1 presenta las conclusiones obtenidas del trabajo realizado. En la sección 5.2 se proponen futuras líneas de trabajo e investigación.

5.1. Conclusiones

En este proyecto se ha trabajado sobre una implementación del PCE para dar soporte a las siguientes soluciones multi-dominio: *Per-domain Path Computation*, *Backward-Recursive PCE-based Computation (BRPC)* y *PCE Jerárquico (H-PCE)*. En particular, se han implementado las siguientes RFCs o drafts:

- J.P. Vasseur (Ed.), A. Ayyangar (Ed.) and R. Zhang, “A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs),” *IETF RFC 5152*, pp. 1-21, February 2008.
- J.P. Vasseur (Ed.), R. Zhang and N. Bitar, J.L. Le Roux, “A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths,” *IETF RFC 5441*, pp. 1-18, April 2009.
- D. King (Ed.) and A. Farrel (Ed.), “The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS,” *IETF Internet-Draft*, pp. 1-28 April 2011.
- F. Zhang, Q. Zhao, O. Gonzalez de Dios, R. Casellas and D. King, “Extensions to Path Computation Element Communication Protocol (PCEP) for Hierarchical Path Computation Elements (PCE),” *IETF Internet-Draft*, pp. 1-14, April 2011.

Además, se han definido escenarios para validar su funcionamiento y comparar las distintas soluciones implementadas. En concreto, se ha caracterizado una topología de red real (*Red Troncal China*) para realizar diferentes experimentos y comparar los protocolos multi-dominio en términos de probabilidad de bloqueo. Con el objetivo de contrastar los resultados se ha utilizado también una topología experimental, cuya información se puede encontrar en el anexo A.

Los resultados muestran que aparece un tipo de bloqueo denominado “robo de longitud de onda” cuando los PCCs intentan reservar los recursos asignados. Esto está condicionado por el tiempo de cómputo de los caminos, es decir, cuanto mayor es este tiempo, existe una mayor probabilidad de que se produzca dicho bloqueo al existir diferente información en la TED del PCE que el estado real de la red. El tiempo de cómputo depende del factor K , que equivale al número de caminos multi-dominio que se comprueban desde el origen hasta el destino. Los mejores resultados en términos de probabilidad de bloqueo se muestran, por tanto, cuando $K = K_{max}$ a costa de aumentar dicho tiempo de cómputo.

Con el objetivo de reducir o incluso eliminar el bloqueo por “robo de longitud de onda” en la reserva de recursos, se han propuesto dos soluciones alternativas: la utilización del algoritmo Round-Robin First-Fit en la elección de longitudes de onda y el uso de un mecanismo de pre-reserva de los recursos del camino. La primera de ellas reduce dicho bloqueo pero no lo elimina completamente. Por tanto, también

reduce el bloqueo total. La segunda elimina dicho bloqueo por completo si se utiliza un temporizador (T_{res}) que pre-reserve los recursos el tiempo suficiente como para asegurar la reserva de los mismos por parte del PCC. Para la implementación de este mecanismo de reserva temporal de los recursos se ha utilizado el siguiente draft:

- O. Gonzalez de Dios, R. Casellas and F. Jimenez Chico, "PCEP Extensions for Temporary Reservation of Computed Path Resources and Support for Limited Context State in PCE," *IETF Internet-Draft*, pp. 1-18, October 2010.

Las pruebas realizadas muestran que, tanto para el BRPC como para el H-PCE, la probabilidad de bloqueo total disminuye con la utilización de las dos mejoras introducidas, mostrándose los mejores resultados con el mecanismo de pre-reserva.

Al realizar una comparativa entre las tres soluciones con la pre-reserva de recursos, se observa que el H-PCE es el que presenta menor probabilidad de bloqueo incluso sin utilizar el factor K máximo, seguido sucesivamente por el BRPC y el per-domain. Para un tráfico ofrecido de $15\ Erlangs$ por nodo, la probabilidad de bloqueo para el per-domain supera el 20%, para el BRPC y $K = K_{max} = 4$ su valor se acerca al 10% y en el caso del H-PCE se tiene para $K = 4$ una probabilidad de casi el 3% y para $K = K_{max} = 20$ del 1%. Por tanto, se puede concluir que el H-PCE presenta las mejores prestaciones en términos de probabilidad de bloqueo, pero debe remarcarse que a medida que el factor K aumenta el retardo computacional en los PCEs también crece.

Fruto de este PFC se ha escrito el artículo "Utilization of Temporary Reservation of Path Computed Resources for Multi-Domain PCE Protocols in WDM Networks", que trata el problema del bloqueo por "robo de longitud de onda" y expone las soluciones implementadas en este PFC haciendo énfasis en la pre-reserva de recursos. Tanto el artículo como la documentación acreditativa se pueden encontrar en el anexo E.

5.2. Futuras líneas de trabajo

Se proponen los siguientes puntos como continuación a este trabajo:

- Inclusión del retardo del plano de control en la reserva de recursos y propagación de la información, propio de una red de comunicaciones actual que demuestre la importancia de un mecanismo de pre-reserva.
- Estudio del impacto del temporizador del mecanismo de pre-reserva con valores que no garanticen la eliminación total del bloqueo por “robo de longitud de onda”.
- Inclusión del Round-Trip Time (RTT) en las comunicaciones PCC-PCE y PCE-PCE.
- Implementación del sistema en un escenario real y estudio del impacto de los elementos del plano de control que intervienen en el proceso.
- Inclusión de un mecanismo de reenvío de solicitudes en caso de recibirse respuestas negativas. Así mismo, reoptimización de los caminos con el objetivo de obtener el máximo provecho de la red.

Referencias

- [1] J. Lobo and F. Jimenez, "Impact of GMPLS on an Integrated Operator," in *WGN5: V Workshop in G/MPLS networks*, pp. 125-128, March 2006.
- [2] J. Gabeiras, V. López, J. Aracil, J. Fernández-Palacios, C. García Argos, O. González de Dios, F. Jiménez Chico and J. Hernández, "Is Multi-Layer Networking Feasible?," *Optical Switching and Networking*, vol. 6, no. 2, pp. 129-140, 2009.
- [3] S. Dasgupta, J. C. de Oliveira and J. P. Vasseur, "Path-Computation-Element-Based Architecture for Interdomain MPLS/GMPLS Traffic Engineering: Overview and Performance," in *IEEE Network*, vol. 21, no. 4, pp. 38-45, July 2007.
- [4] ITU-T, "Architecture for the automatically switched optical network (ASON)," G.8080/Y.1304 *Recommendation*, February 2005. Online: <http://www.itu.int/rec/T-REC-G.8080/en>
- [5] E. Mannie (Ed.), "Generalized Multi-Protocol Label Switching (GMPLS) Architecture," *IETF RFC 3945*, pp. 1-69, October 2004. Online: <http://tools.ietf.org/html/rfc3945>
- [6] D. Awduche, L. Berger, D. Gan and T. Li, V. Srinivasan, G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels," *IETF RFC 3209*, pp. 1-61, December 2001. Online: <http://tools.ietf.org/html/rfc3209>
- [7] L. Berger (Ed.), "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol -Traffic Engineering (RSVP-TE) Extensions," *IETF RFC 3473*, pp. 1-42, January 2003. Online: <http://tools.ietf.org/html/rfc3473>
- [8] A. Farrel, J.P. Vasseur and J. Ash, "A Path Computation Element (PCE)-Based Architecture," *IETF RFC 4655*, pp. 1-40, August 2006. Online: <http://tools.ietf.org/html/rfc4655>
- [9] V. López, B. Huiszoon, J. Fernández-Palacios, O. González de Dios and J. Aracil, "Path Computation Element in Telecom Networks: Recent Developments and Standardization Activities," in *Optical Networking Design and Modelling (ONDM)*, February 2010.
- [10] J.L. Le Roux (Ed.), "Requirements for Path Computation Element (PCE) Discovery," *IETF RFC 4674*, pp. 1-19, October 2006. Online: <http://tools.ietf.org/html/rfc4674>
- [11] J.L. Le roux (Ed.), J.P. Vasseur (Ed.), Y. Ikejiri and R. Zhang, "OSPF Protocol Extensions for Path Computation Element (PCE) Discovery," *IETF RFC 5088*, pp. 1-20, January 2008. Online: <http://tools.ietf.org/html/rfc5088>
- [12] J.L. Le roux (Ed.), J.P. Vasseur (Ed.), Y. Ikejiri and R. Zhang, "IS-IS Protocol Extensions for Path Computation Element (PCE) Discovery," *IETF RFC 5089*, pp. 1-17, January 2008. Online: <http://tools.ietf.org/html/rfc5089>
- [13] A. Lindem (Ed.), N. Shen, J.P. Vasseur, R. Aggarwal and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities," *IETF RFC 4970*, pp. 1-13, July 2007. Online: <http://tools.ietf.org/html/rfc4970>

- [14] J.P. Vasseur (Ed.), N. Shen (Ed.) and R. Aggarwal (Ed.), "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information," *IETF RFC 4971*, pp. 1-9, July 2007. Online: <http://tools.ietf.org/html/rfc4971>
- [15] J.P. Vasseur (Ed.) and J.L. Le Roux (Ed.), "Path Computation Element (PCE) Communication Protocol (PCEP)," *IETF RFC 5440*, pp. 1-87, March 2009. Online: <http://tools.ietf.org/html/rfc5440>
- [16] CIDR report, June 2011. Online: <http://www.cidr-report.org/>
- [17] A. Farrel, J.P. Vasseur and A. Ayyangar, "A Framework for the Inter-Domain Multiprotocol Label Switching Traffic Engineering," *IETF RFC 4726*, pp. 1-22, November 2006. Online: <http://tools.ietf.org/html/rfc4726>
- [18] Y. Rekhter (Ed.) and T. Li (Ed.), "A Border Gateway Protocol 4 (BGP-4)," *IETF RFC 1771*, pp. 1-57, March 1995. Online: <http://tools.ietf.org/html/rfc1771>
- [19] S. Okamoto, H. Otsuki and T. Otani, "Multi-ASON and GMPLS network domain interworking challenges," *IEEE Communications Magazine*, vol. 46, no. 6, pp. 88-93, 2008.
- [20] ITU-T, Rec. G.7713.2, "Distributed Call and Connection Management: Signaling Mechanism Using GMPLS RSVP-TE," March 2003.
- [21] J.P. Vasseur (Ed.), A. Ayyangar (Ed.) and R. Zhang, "A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)," *IETF RFC 5152*, pp. 1-21, February 2008. Online: <http://tools.ietf.org/html/rfc5152>
- [22] F. Aslam, Z.A. Uzmi, and A. Farrel, "Interdomain Path Computation: Challenges and Solutions for Label Switched Networks," *IEEE Communications Magazine*, vol. 45, no. 10, pp. 94-101, 2007.
- [23] D. King (Ed.) and A. Farrel (Ed.), "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS," *IETF Internet-Draft*, pp. 1-28 April 2011. Online: <http://tools.ietf.org/search/draft-king-pce-hierarchy-fwk-06>
- [24] J.P. Vasseur (Ed.), R. Zhang and N. Bitar, J.L. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths," *IETF RFC 5441*, pp. 1-18, April 2009. Online: <http://tools.ietf.org/html/rfc5441>
- [25] M. Chen, R. Zhang and X. Duan, "IS-IS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering," *IETF RFC 5316*, pp. 1-19, December 2008. Online: <http://tools.ietf.org/html/rfc5316>
- [26] M. Chen, R. Zhang and X. Duan, "OSPF Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering," *IETF RFC 5392*, pp. 1-17, January 2009. Online: <http://tools.ietf.org/html/rfc5392>
- [27] J.L. Le roux, J.P. Vasseur and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communications Protocol (PCEP)," *IETF RFC 5541*, pp. 1-23, June 2009. Online: <http://tools.ietf.org/html/rfc5541>

- [28] I. Bryskin, D. Papadimitriou, L. Berger and J. Ash, "Policy-Enabled Path Computation Framework," *IETF RFC 5394*, pp.1-36, December 2008. Online: <http://tools.ietf.org/html/rfc5394>
- [29] R. Bradford (Ed.), J.P. Vasseur and A. Farrel, "Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Path-Key-Based Mechanism," *IETF RFC 5520*, pp. 1-19, April 2009. Online: <http://tools.ietf.org/html/rfc5520>
- [30] F. Zhang, Q. Zhao, O. Gonzalez de Dios, R. Casellas and D. King, "Extensions to Path Computation Element Communication Protocol (PCEP) for Hierarchical Path Computation Elements (PCE)," *IETF Internet-Draft*, pp. 1-14, April 2011. Online: <http://tools.ietf.org/html/draft-zhang-pce-hierarchy-extensions-00>
- [31] Y. Zhao, J. Zhang, Y. Ji and W. Gu, "Routing and Wavelength Assignment Problem in PCE-Based Wavelength-Switched Optical Networks," *Journal of Optical Communications and Networking*, vol. 2, no. 4, pp. 196-205, 2010.
- [32] ITU-T, "Physical transfer functions of optical network elements," *G.680 Recommendation*, pp. 1-68, February 2005. Online: <http://www.itu.int/rec/T-REC-G.680/en>
- [33] J. Strand, A.L. Chiu and R. Tkach, "Issues for routing in the optical layer," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 81-87, February 2001.
- [34] L. Berger (Ed.), "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description," *IETF RFC 3471*, pp. 1-34, January 2003. Online: <http://tools.ietf.org/html/rfc3471>
- [35] O. Gonzalez de Dios, R. Casellas and F. Jimenez Chico, "PCEP Extensions for Temporary Reservation of Computed Path Resources and Support for Limited Context State in PCE," *IETF Internet-Draft*, pp. 1-18, October 2010. Online: <http://tools.ietf.org/html/draft-gonzalezdedios-pce-reservation-state-00>
- [36] J.P. Vasseur (Ed.), J.L. Le Roux and Y. Ikejiri, "A Set of Monitoring Tools for Path Computation Element (PCE)-Based Architecture," *IETF RFC 5886*, pp. 1-26, June 2010. Online: <http://tools.ietf.org/html/rfc5886>

Anexo A. Validación de resultados

Con el objetivo de evaluar las distintas soluciones implementadas y contrastar los resultados en términos de probabilidad de bloqueo y observar el efecto del bloqueo por “robo de longitud de onda”, se han realizado los mismos experimentos estudiados en las secciones 4.2 y 4.3 con una topología diferente. Dicha topología (Figura A.1) está formada por cuatro ASes, todos ellos con seis nodos de red. Los cuatro ASes tienen la misma topología intra-dominio en forma de ocho. En la Tabla A.1 se pueden observar las características de esta topología.

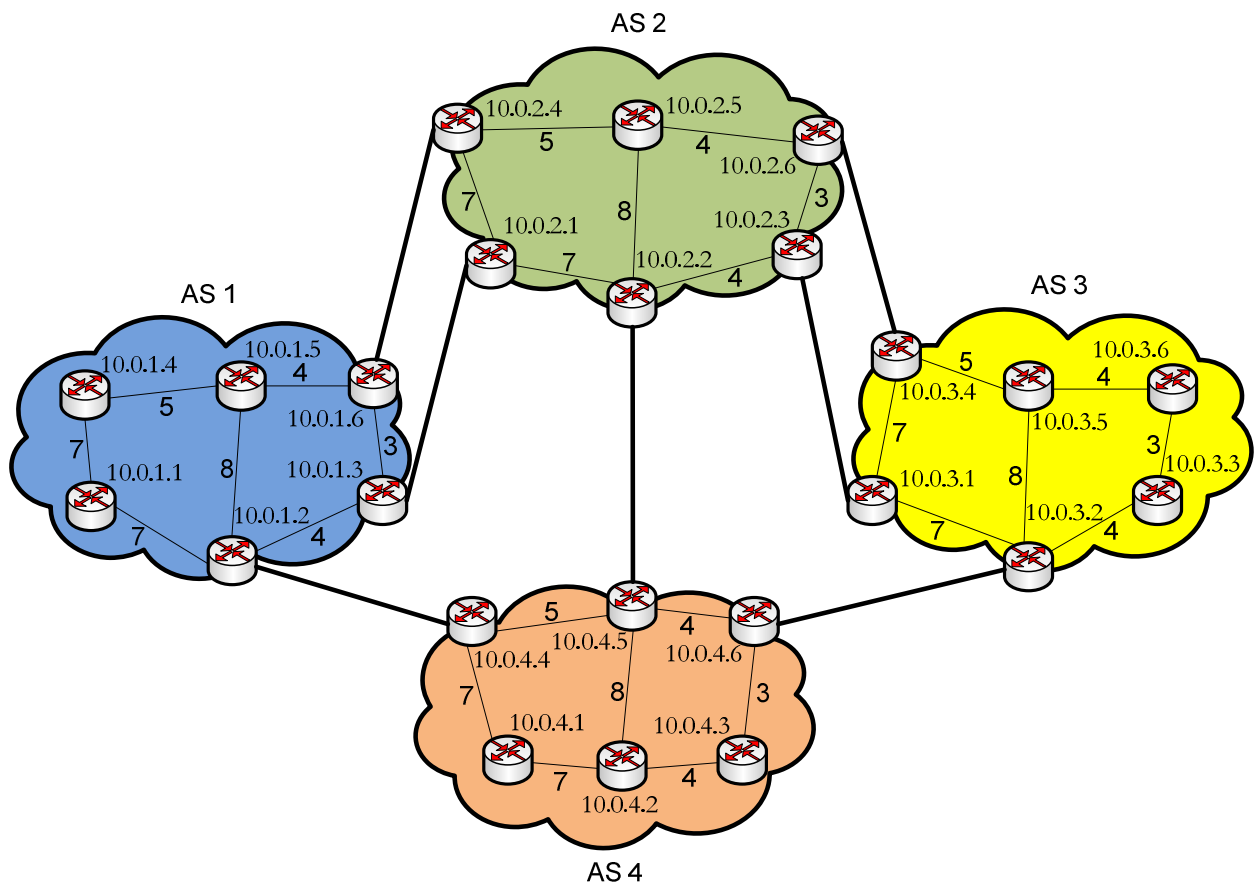


Figura A.1 - Topología experimental *Ochos*

Tabla A.1 – Características de la topología experimental *Ochos* (24 nodos)

Número de AS	Número de nodos	Número de enlaces (bidireccionales)	Grado medio de los nodos
1, 2, 3, 4	6	7	2.33

En este caso, el valor máximo de K para el BRPC es de $K_{max} = 4$ y para el H-PCE de $K_{max} = 10$.

A continuación se muestran los resultados de los experimentos realizados:

- Estabilidad para per-domain

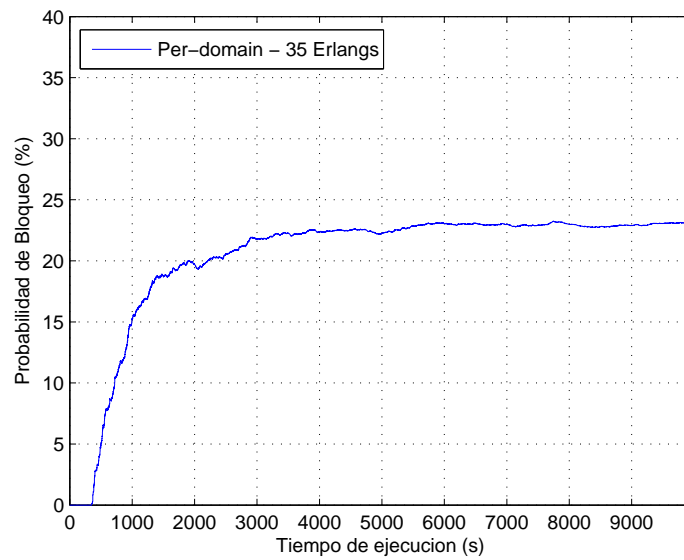


Figura A.2 - Estabilidad para per-domain

- Estabilidad para BRPC

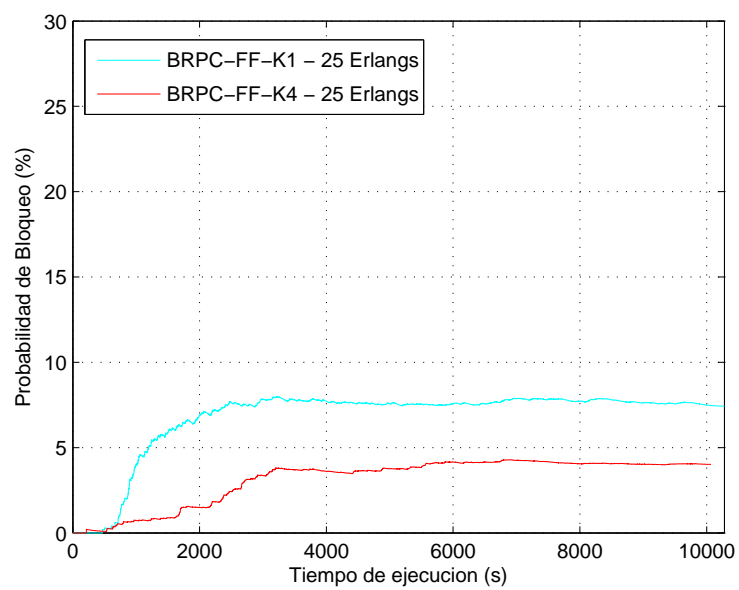


Figura A.3 - Estabilidad para BRPC

- Estabilidad para H-PCE

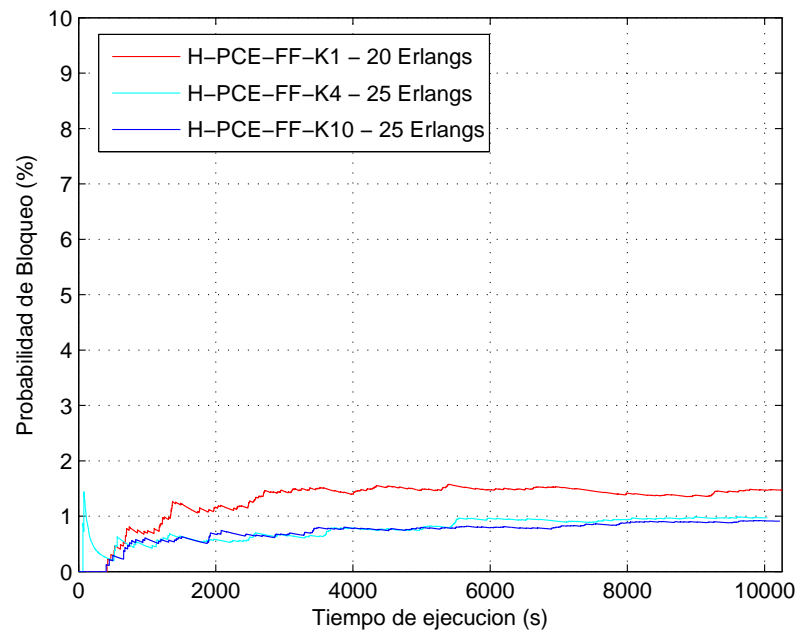


Figura A.4 - Estabilidad para H-PCE

- Probabilidad de bloqueo para BRPC variando K con el algoritmo FF:

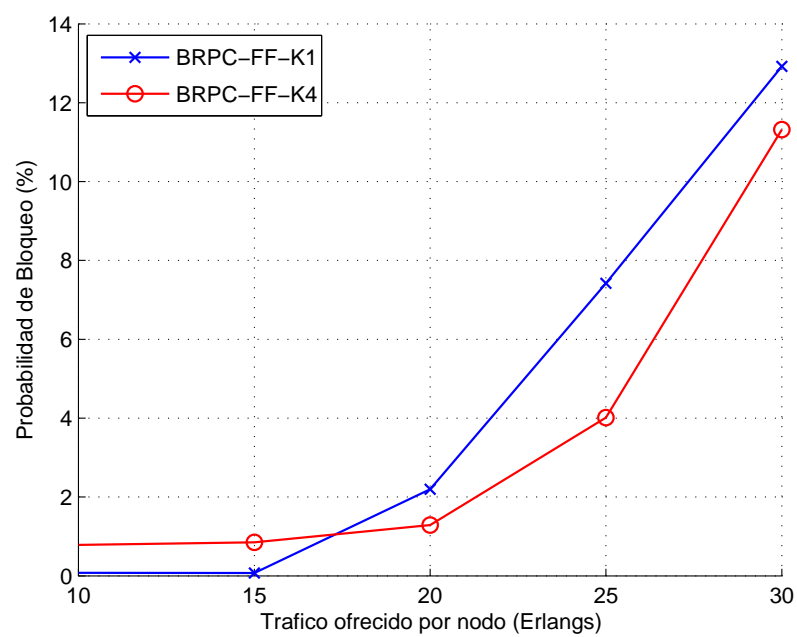


Figura A.5 - Probabilidad de bloqueo para BRPC variando K con el algoritmo FF

- Tipos de bloqueo para BRPC variando K con el algoritmo FF:

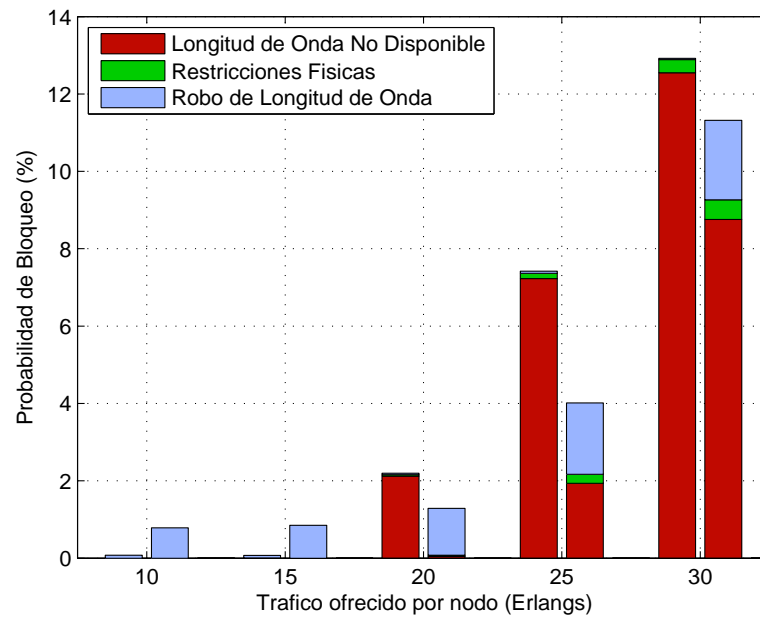


Figura A.6 - Tipos de bloqueo para BRPC variando K con el algoritmo FF

- Probabilidad de bloqueo para H-PCE variando K con el algoritmo FF:

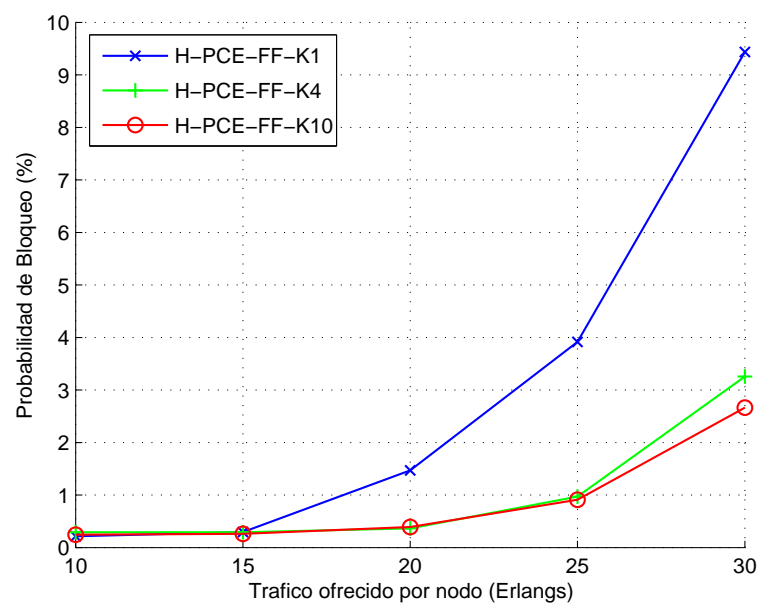


Figura A.7 - Probabilidad de bloqueo para H-PCE variando K con el algoritmo FF

- Tipos de bloqueo para H-PCE variando K con el algoritmo FF:

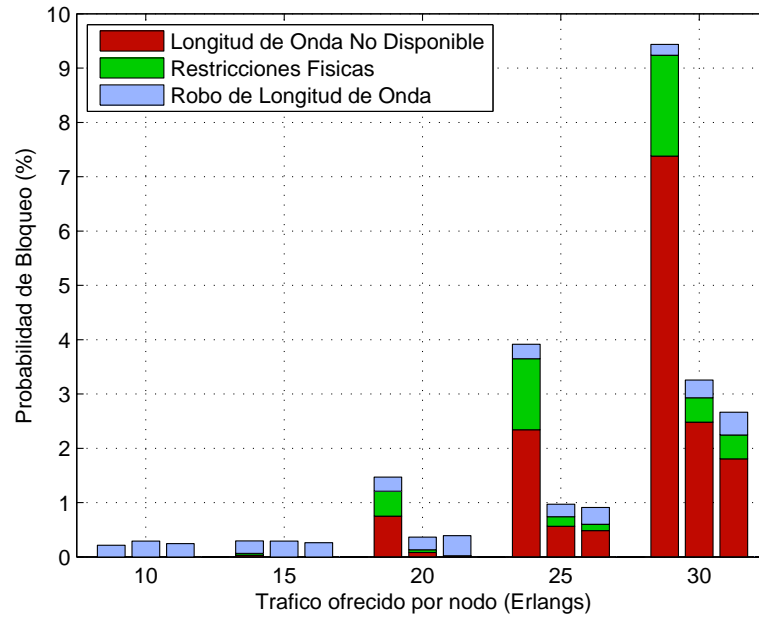


Figura A.8 - Tipos de bloqueo para H-PCE variando K con el algoritmo FF

- Probabilidad de bloqueo para per-domain con el algoritmo FF:

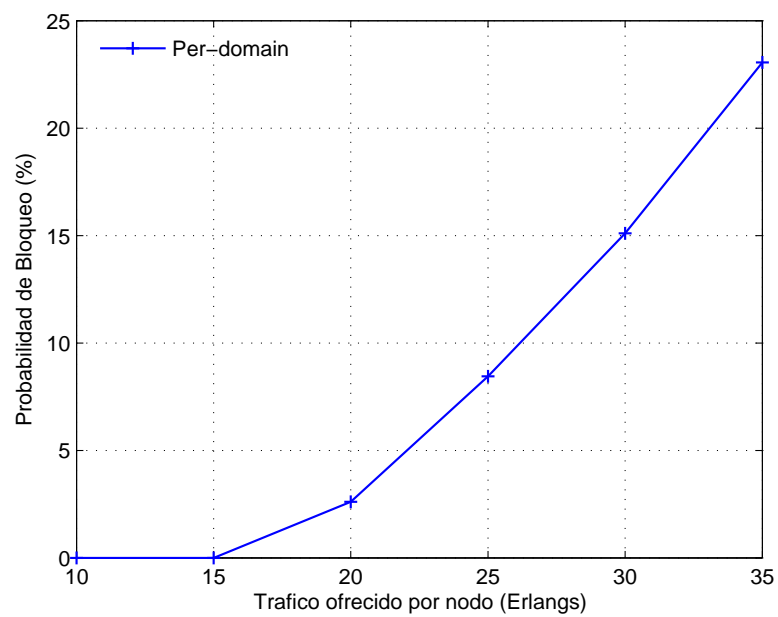


Figura A.9 - Probabilidad de bloqueo para per-domain con el algoritmo FF

- Probabilidad de bloqueo para BRPC y $K = K_{max}$ con los algoritmos FF, RR-FF y PR:

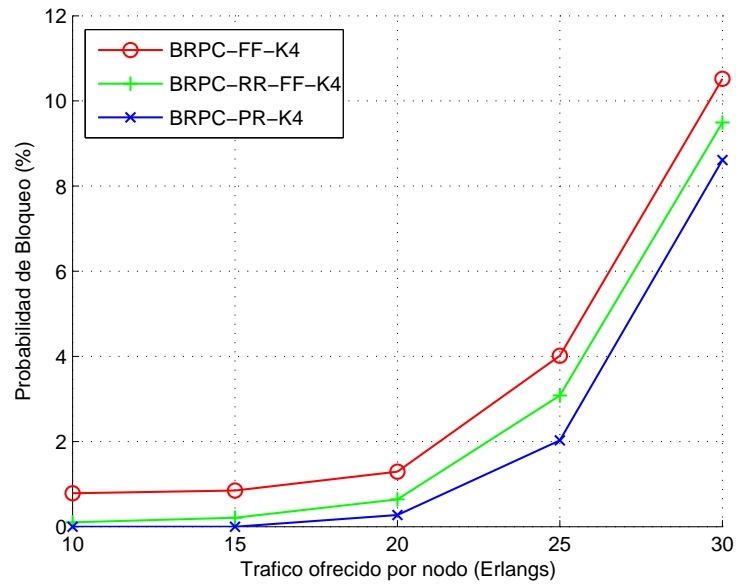


Figura A.10 - Probabilidad de bloqueo para BRPC con los algoritmos FF y RR-FF y PR

- Tipos de bloqueo para BRPC y $K = K_{max}$ con los algoritmos FF, RR-FF y PR:

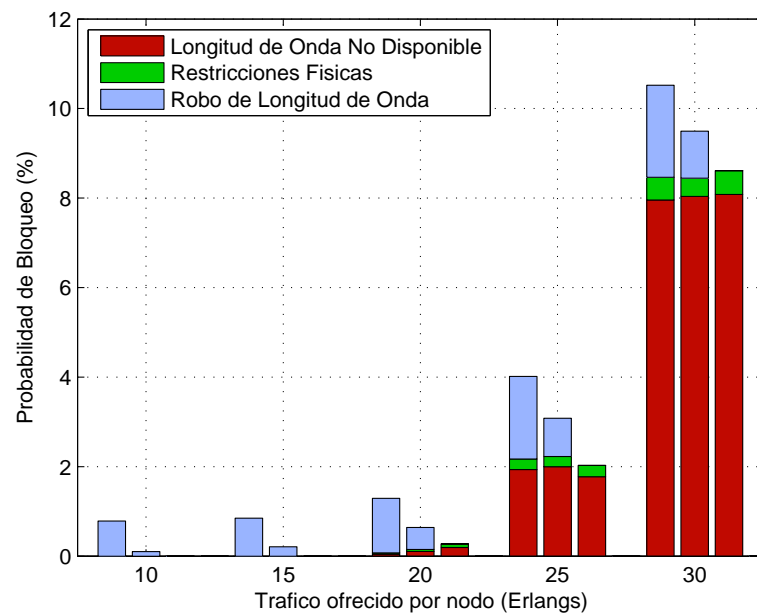


Figura A.11 - Tipos de bloqueo para BRPC con los algoritmos FF y RR-FF y PR

- Probabilidad de bloqueo para H-PCE y $K = K_{max}$ con los algoritmos FF, RR-FF y PR:

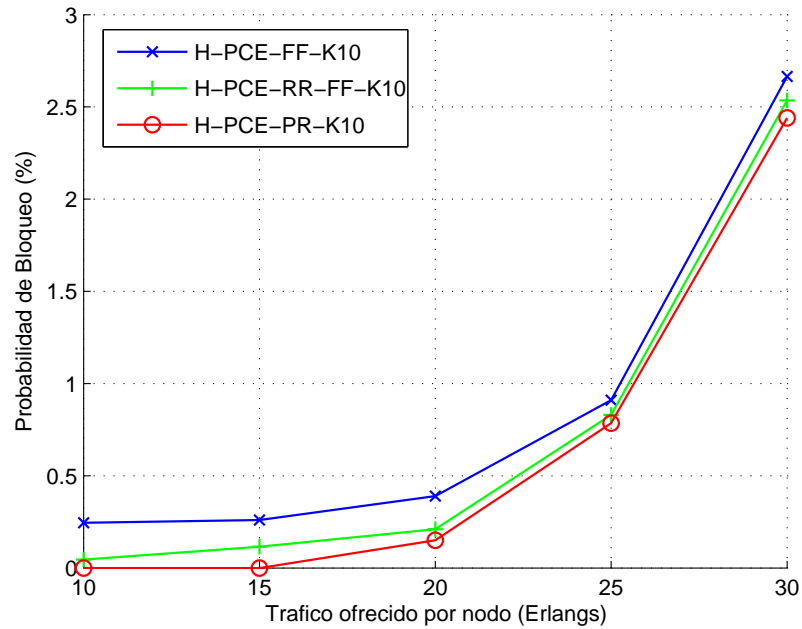


Figura A.12 - Probabilidad de bloqueo para H-PCE con los algoritmos FF y RR-FF y PR

- Tipos de bloqueo para H-PCE y $K = K_{max}$ con los algoritmos FF, RR-FF y PR:

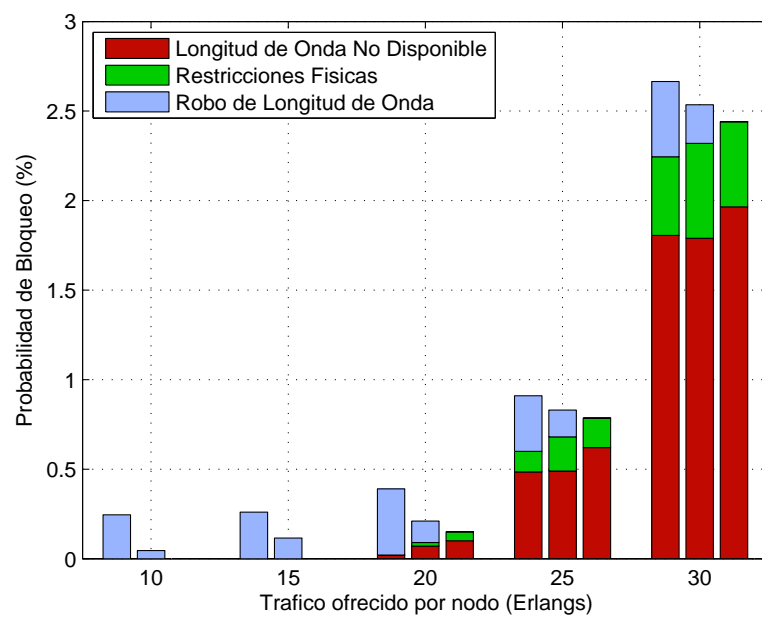


Figura A.13 - Tipos de bloqueo para H-PCE con los algoritmos FF y RR-FF y PR

- Comparativa en términos de probabilidad de bloqueo para BRPC, H-PCE y per-domain para distintos valores de K con el mecanismo de pre-reserva:

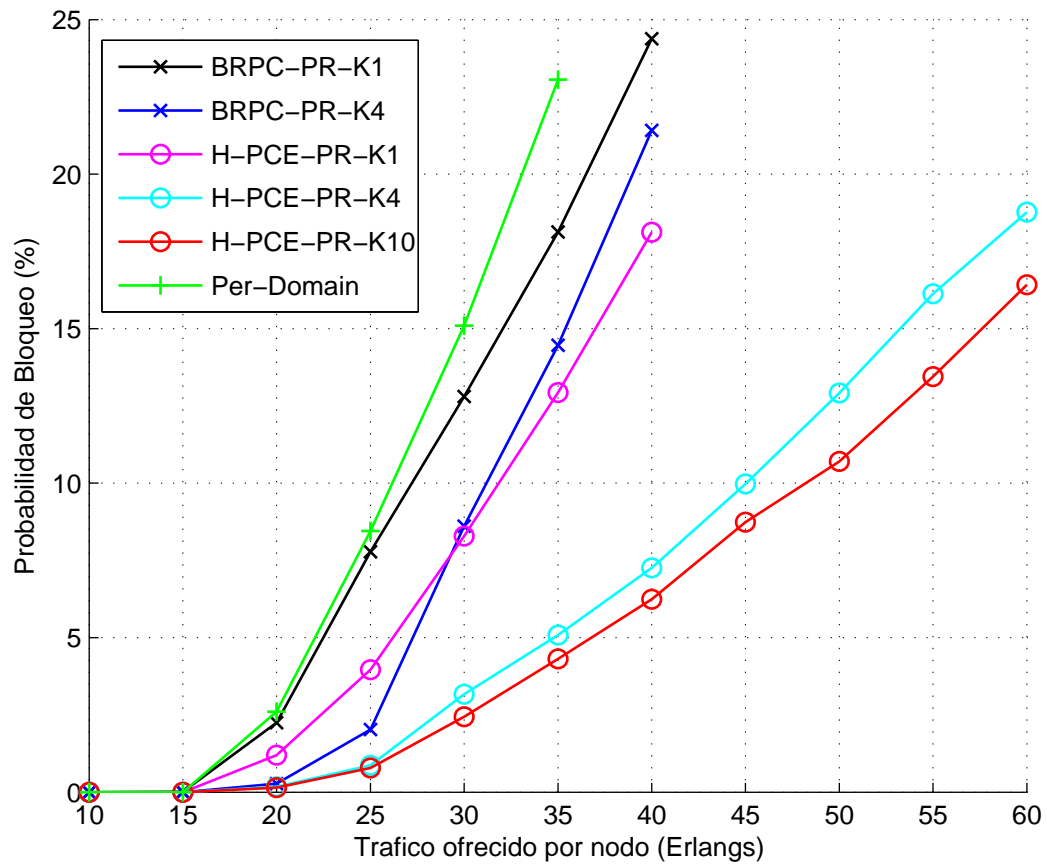


Figura A.14 - Comparativa en términos de probabilidad de bloqueo con el mecanismo PR

Anexo B. Manual de usuario

Información previa

Los distintos escenarios se encuentran divididos en función de los tres protocolos implementados (per-domain, BRPC y H-PCE), así como de los algoritmos First-Fit (FF) y Round-Robin First-Fit (RR-FF) y del mecanismo de Pre-Reserva (PR). Así, en el mismo directorio se pueden encontrar, por separado, los siguientes escenarios:

- **PDPC_FF**: Protocolo per-domain con algoritmo FF.
- **BRPC_FF**: Protocolo BRPC con algoritmo FF.
- **BRPC_RR-FF**: Protocolo BRPC con algoritmo RR-FF.
- **BRPC_PR**: Protocolo BRPC con mecanismo de PR.
- **H-PCE_FF**: Protocolo H-PCE con algoritmo FF.
- **H-PCE_RR-FF**: Protocolo H-PCE con algoritmo RR-FF.
- **H-PCE_PR**: Protocolo H-PCE con mecanismo de PR.

Todos ellos contienen los mismos ficheros de código, además de una carpeta llamada *topologies* con toda la información referente a la topología de red: dominios, conexiones, nodos de borde, costes, etc.

Existen dos variables importantes a la hora de realizar los experimentos: el factor K y el número de peticiones a realizar por un PCC Fuente. En cada escenario puede encontrarse un fichero *léeme.txt* con información útil acerca de estas dos variables.

Compilación

El programa cuenta con un fichero Makefile para compilarlo mediante el comando *make*. Una vez compilado, se generan los siguientes ficheros ejecutables:

- **PCEServer**: Se trata del PCE. Por cada PCE del escenario el usuario deberá ejecutar uno distinto.
- **PCC**: Es el “módulo PCC” de los PCEs. Su ejecución forma parte del proceso de comunicación entre PCEs, por lo que el usuario no deberá ejecutarlo.
- **PCCIni**: Es el PCC Fuente. El usuario deberá ejecutar uno por cada dominio de la topología de red.
- **TruncateDB**: Es un ejecutable que sirve para reiniciar los valores de las tablas de la base de datos. Antes de cada simulación el usuario deberá ejecutarlo.

Ejecución

Para esta fase se seguirá el ejemplo de la *Red Troncal China* [31] que, para mayor comodidad, se vuelve reproducir en la Figura B.1.

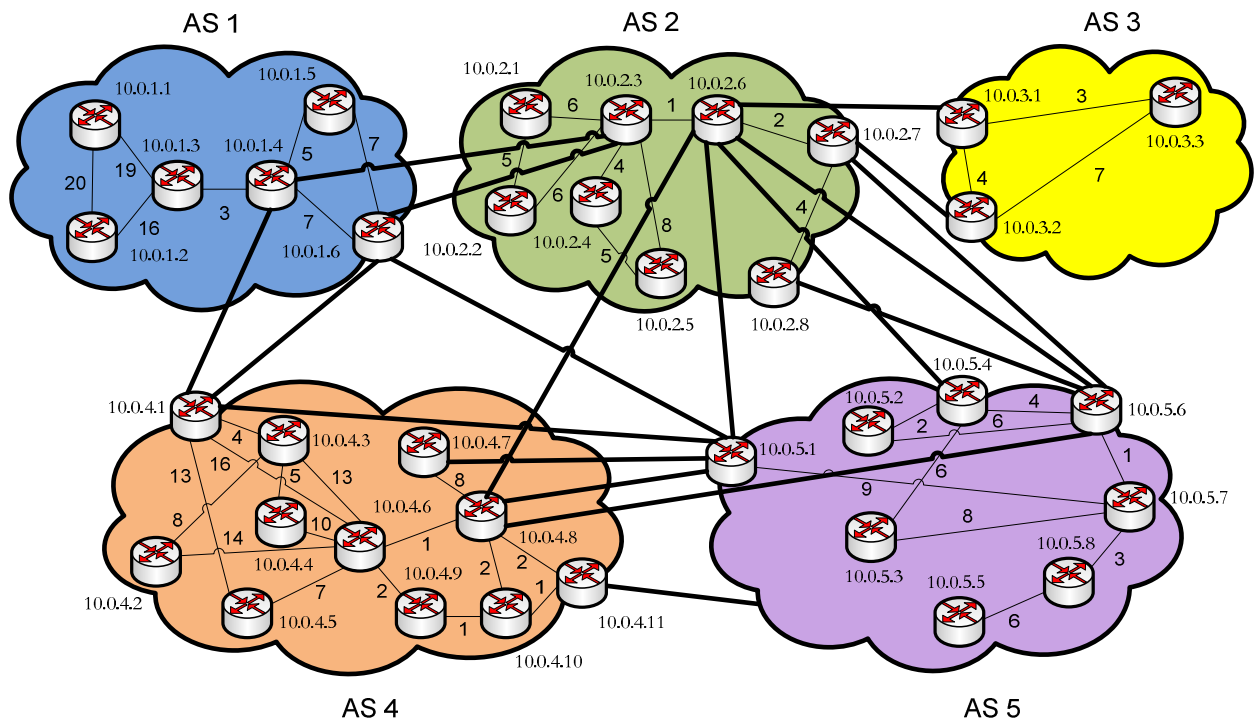


Figura B.1 - *Red Troncal China* [31]

En cuanto a los protocolos, se deben diferenciar dos casos: por un lado el per-domain y el BRPC y, por otro, el H-PCE.

- Per-domain y BRPC: Se deberán ejecutar 5 PCCs Fuente y 5 PCEs (uno por cada dominio).
- H-PCE: Se deberán ejecutar 5 PCCs Fuente (uno por cada dominio) y 6 PCEs (uno por cada dominio más el PCE padre). En este caso, el PCE padre sería el PCE 6.

Para ejecutar un PCE se deberá introducir un único argumento de entrada, correspondiente al puerto TCP donde se realizará la conexión. Al estar trabajando en la interfaz de *loopback*, se han elegido los puertos *400X*, en donde *X* es el número de dominio al que corresponde el PCE. De esta manera, se introducirá la siguiente línea:

```
./PCEServer 400X
```


Como ejemplo, véase la Figura B.2, en donde se está ejecutando el PCE 1. Tras introducir la línea de entrada, el PCE tendrá el aspecto de la Figura B.2 (abajo), indicando que está activo y preparado recibir solicitudes.

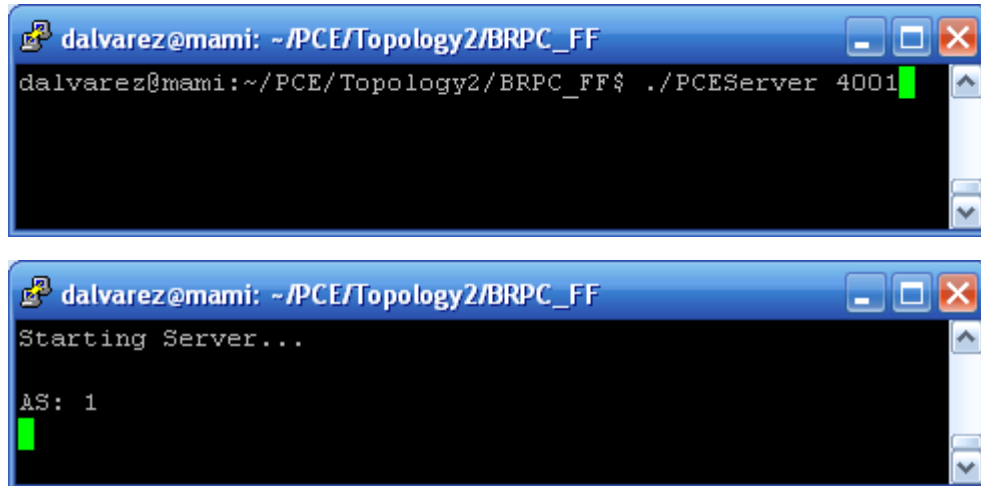
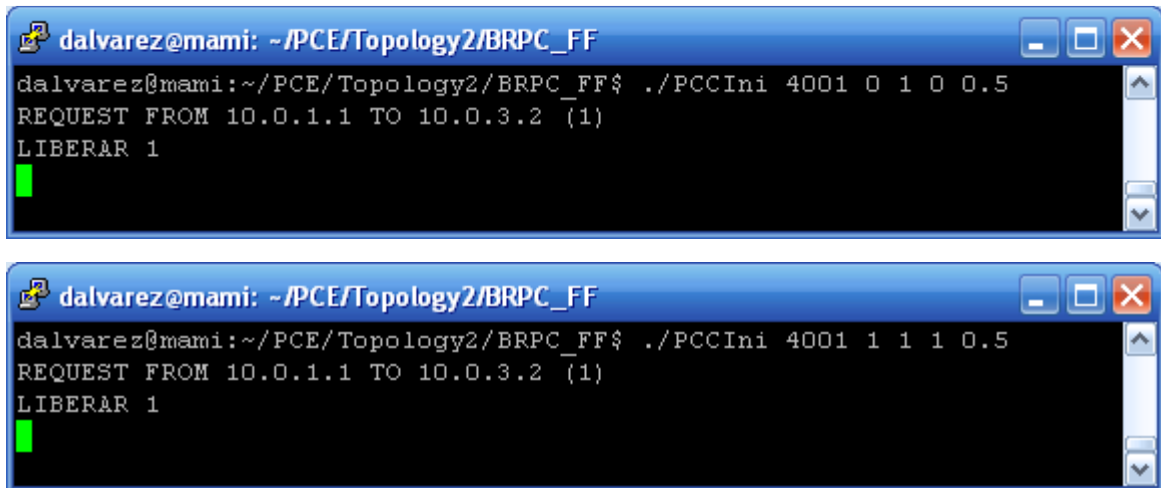


Figura B.2 - Ejecución del PCE 1 antes (arriba) y después (abajo)

En cuanto a los PCCs Fuente, se explican a continuación los argumentos de entrada correspondientes:

- 1) **Puerto TCP:** Con el mismo formato que para el PCE del mismo dominio (400X).
- 2) **Tasa media de llegadas de peticiones:** Su valor es 0 para una tasa determinista y 1 para una tasa exponencial.
- 3) **Valor de tasa media de llegadas de peticiones:** Correspondiente al número de nodos que contiene cada dominio.
- 4) **Tasa media de servicio de peticiones:** Su valor es 0 para una tasa determinista y 1 para una tasa exponencial.
- 5) **Valor de la tasa media de servicio de peticiones:** Correspondiente al tiempo que los caminos estarán establecidos.

Como ejemplo puede observarse la Figura B.3, en donde se desea ejecutar el PCC Fuente del dominio 1.



The image shows two terminal windows. The top window shows the command `./PCCIni 4001 0 1 0 0.5` and the output `REQUEST FROM 10.0.1.1 TO 10.0.3.2 (1)` and `LIBERAR 1`. The bottom window shows the command `./PCCIni 4001 1 1 1 0.5` and the same output.

```

dalvarez@mami: ~/PCE/Topology2/BRPC_FF
dalvarez@mami:~/PCE/Topology2/BRPC_FF$ ./PCCIni 4001 0 1 0 0.5
REQUEST FROM 10.0.1.1 TO 10.0.3.2 (1)
LIBERAR 1

dalvarez@mami: ~/PCE/Topology2/BRPC_FF
dalvarez@mami:~/PCE/Topology2/BRPC_FF$ ./PCCIni 4001 1 1 1 0.5
REQUEST FROM 10.0.1.1 TO 10.0.3.2 (1)
LIBERAR 1

```

Figura B.3 - Ejecución del PCC Fuente 1 con tasas deterministas (arriba) y exponenciales (abajo)

Resultados

Los datos acerca del estado de las peticiones se recogen en distintos ficheros de texto. A continuación se presenta una lista con información sobre dichos ficheros:

- ***establishment_info.txt***: Por cada petición se presenta una línea con información sobre el identificador de petición, si ha habido bloqueo (indicando el motivo) o no lo ha habido (indicando el número de saltos del camino) y, por último, el tiempo en formato Unix en que ha llegado la respuesta a la petición. Los motivos del bloqueo son tres: *No Wavelength* indica que no había longitudes de onda disponibles para asignar, *Physical Restrictions* que el bloqueo por restricciones físicas de los enlaces y *Wavelength not available* que se ha producido bloqueo por “robo de longitud de onda”. Un ejemplo de petición satisfactoria es

ID: 1 (OK) Hops: 3. Time: 1305628562.401617

La misma petición podría dar lugar a un bloqueo y se presentaría como en el siguiente ejemplo

ID: 1 (NOK) No Wavelength. Time: 1305628562.401617

- ***time_arrival.txt***: Recoge información en dos columnas. La primera es el identificador de petición y la segunda el tiempo (en formato Unix) en que se ha generado la petición.
- ***time_reply.txt***: En una sola columna, se recogen los tiempos (en microsegundos) que tardan las peticiones en ser atendidas, es decir, desde que se envía la petición al PCE hasta que llega la respuesta.

- ***time_release.txt***. En formato de dos columnas, la primera de ellas es el identificador de petición y la segunda el tiempo (en formato Unix) en que los recursos asignados al camino han sido liberados.
- ***time_ksp.txt***. En una columna se recogen los tiempos de cómputo (en microsegundos) de los algoritmos de enrutamiento, validación de restricciones y elección de longitud de onda.

Anexo C. Presupuesto

1. Ejecución Material

- Compra de ordenador personal (Software incluido)..... 3.000 €
 - Alquiler de servidor durante un año1.000 €
 - Material de oficina50 €
- Total de ejecución material..... 4.050 €

2) Gastos generales

- 16 % sobre Ejecución Material..... 648 €

3) Beneficio Industrial

- 6 % sobre Ejecución Material..... 243 €

4) Honorarios Proyecto

- 1.600 horas a 18 € / hora 28.800 €

5) Material fungible

- Gastos de impresión 80 €
- Encuadernación 20 €

6) Subtotal del presupuesto

- Subtotal Presupuesto..... 33.841 €

7) I.V.A. aplicable

- 18% Subtotal Presupuesto 6.091,38 €

8) Total presupuesto

- Total Presupuesto..... 39.932,38 €

Madrid, Septiembre de 2011

El Ingeniero Jefe de Proyecto

Fdo.: Diego Álvarez Martínez
Ingeniero Superior de Telecomunicación

Anexo D. Pliego de condiciones

Este documento contiene las condiciones legales que guiarán la realización de este proyecto: *Implementación y comparación de soluciones basadas en Path Computation Element (PCE) para entornos multi-dominio*. En lo que sigue, se supondrá que el proyecto ha sido encargado por una empresa cliente a una empresa consultora con la finalidad de realizar dicho sistema. Dicha empresa ha debido desarrollar una línea de investigación con objeto de elaborar el proyecto. Esta línea de investigación, junto con el posterior desarrollo de los programas está amparada por las condiciones particulares del siguiente pliego.

Supuesto que la utilización industrial de los métodos recogidos en el presente proyecto ha sido decidida por parte de la empresa cliente o de otras, la obra a realizar se regulará por las siguientes:

Condiciones generales

1. La modalidad de contratación será el concurso. La adjudicación se hará, por tanto, a la proposición más favorable sin atender exclusivamente al valor económico, dependiendo de las mayores garantías ofrecidas. La empresa que somete el proyecto a concurso se reserva el derecho a declararlo desierto.

2. El montaje y mecanización completa de los equipos que intervengan será realizado totalmente por la empresa licitadora.

3. En la oferta, se hará constar el precio total por el que se compromete a realizar la obra y el tanto por ciento de baja que supone este precio en relación con un importe límite si este se hubiera fijado.

4. La obra se realizará bajo la dirección técnica de un Ingeniero Superior de Telecomunicación, auxiliado por el número de Ingenieros Técnicos y Programadores que se estime preciso para el desarrollo de la misma.

5. Aparte del Ingeniero Director, el contratista tendrá derecho a contratar al resto del personal, pudiendo ceder esta prerrogativa a favor del Ingeniero Director, quien no estará obligado a aceptarla.

6. El contratista tiene derecho a sacar copias a su costa de los planos, pliego de condiciones y presupuestos. El Ingeniero autor del proyecto autorizará con su firma las copias solicitadas por el contratista después de confrontarlas.

7. Se abonará al contratista la obra que realmente ejecute con sujeción al proyecto que sirvió de base para la contratación, a las modificaciones autorizadas por la superioridad o a las órdenes que con arreglo a sus facultades le hayan comunicado por escrito al Ingeniero Director de obras siempre que dicha

obra se haya ajustado a los preceptos de los pliegos de condiciones, con arreglo a los cuales, se harán las modificaciones y la valoración de las diversas unidades sin que el importe total pueda exceder de los presupuestos aprobados. Por consiguiente, el número de unidades que se consignan en el proyecto o en el presupuesto, no podrá servirle de fundamento para entablar reclamaciones de ninguna clase, salvo en los casos de rescisión.

8. Tanto en las certificaciones de obras como en la liquidación final, se abonarán los trabajos realizados por el contratista a los precios de ejecución material que figuran en el presupuesto para cada unidad de la obra.

9. Si excepcionalmente se hubiera ejecutado algún trabajo que no se ajustase a las condiciones de la contrata pero que sin embargo es admisible a juicio del Ingeniero Director de obras, se dará conocimiento a la Dirección, proponiendo a la vez la rebaja de precios que el Ingeniero estime justa y si la Dirección resolviera aceptar la obra, quedará el contratista obligado a conformarse con la rebaja acordada.

10. Cuando se juzgue necesario emplear materiales o ejecutar obras que no figuren en el presupuesto de la contrata, se evaluará su importe a los precios asignados a otras obras o materiales análogos si los hubiere y cuando no, se discutirán entre el Ingeniero Director y el contratista, sometiéndolos a la aprobación de la Dirección. Los nuevos precios convenidos por uno u otro procedimiento, se sujetarán siempre al establecido en el punto anterior.

11. Cuando el contratista, con autorización del Ingeniero Director de obras, emplee materiales de calidad más elevada o de mayores dimensiones de lo estipulado en el proyecto, o sustituya una clase de fabricación por otra que tenga asignado mayor precio o ejecute con mayores dimensiones cualquier otra parte de las obras, o en general, introduzca en ellas cualquier modificación que sea beneficiosa a juicio del Ingeniero Director de obras, no tendrá derecho sin embargo, sino a lo que le correspondería si hubiera realizado la obra con estricta sujeción a lo proyectado y contratado.

12. Las cantidades calculadas para obras accesorias, aunque figuren por partida alzada en el presupuesto final (general), no serán abonadas sino a los precios de la contrata, según las condiciones de la misma y los proyectos particulares que para ellas se formen, o en su defecto, por lo que resulte de su medición final.

13. El contratista queda obligado a abonar al Ingeniero autor del proyecto y director de obras así como a los Ingenieros Técnicos, el importe de sus respectivos honorarios facultativos por formación del proyecto, dirección técnica y administración en su caso, con arreglo a las tarifas y honorarios vigentes.

14. Concluida la ejecución de la obra, será reconocida por el Ingeniero Director que a tal efecto designe la empresa.

15. La garantía definitiva será del 4% del presupuesto y la provisional del 2%.

16. La forma de pago será por certificaciones mensuales de la obra ejecutada, de acuerdo con los precios del presupuesto, deducida la baja si la hubiera.

17. La fecha de comienzo de las obras será a partir de los 15 días naturales del replanteo oficial de las mismas y la definitiva, al año de haber ejecutado la provisional, procediéndose si no existe reclamación alguna, a la reclamación de la fianza.

18. Si el contratista al efectuar el replanteo, observase algún error en el proyecto, deberá comunicarlo en el plazo de quince días al Ingeniero Director de obras, pues transcurrido ese plazo será responsable de la exactitud del proyecto.

19. El contratista está obligado a designar una persona responsable que se entenderá con el Ingeniero Director de obras, o con el delegado que éste designe, para todo relacionado con ella. Al ser el Ingeniero Director de obras el que interpreta el proyecto, el contratista deberá consultarle cualquier duda que surja en su realización.

20. Durante la realización de la obra, se girarán visitas de inspección por personal facultativo de la empresa cliente, para hacer las comprobaciones que se crean oportunas. Es obligación del contratista, la conservación de la obra ya ejecutada hasta la recepción de la misma, por lo que el deterioro parcial o total de ella, aunque sea por agentes atmosféricos u otras causas, deberá ser reparado o reconstruido por su cuenta.

21. El contratista, deberá realizar la obra en el plazo mencionado a partir de la fecha del contrato, incurriendo en multa, por retraso de la ejecución siempre que éste no sea debido a causas de fuerza mayor. A la terminación de la obra, se hará una recepción provisional previo reconocimiento y examen por la dirección técnica, el depositario de efectos, el interventor y el jefe de servicio o un representante, estampando su conformidad el contratista.

22. Hecha la recepción provisional, se certificará al contratista el resto de la obra, reservándose la administración el importe de los gastos de conservación de la misma hasta su recepción definitiva y la fianza durante el tiempo señalado como plazo de garantía. La recepción definitiva se hará en las mismas condiciones que la provisional, extendiéndose el acta correspondiente. El Director Técnico propondrá a la Junta Económica la devolución de la fianza al contratista de acuerdo con las condiciones económicas legales establecidas.

23. Las tarifas para la determinación de honorarios, reguladas por orden de la Presidencia del Gobierno el 19 de Octubre de 1961, se aplicarán sobre el denominado en la actualidad “Presupuesto de

Ejecución de Contrata” y anteriormente llamado ”Presupuesto de Ejecución Material” que hoy designa otro concepto.

Condiciones particulares

La empresa consultora, que ha desarrollado el presente proyecto, lo entregará a la empresa cliente bajo las condiciones generales ya formuladas, debiendo añadirse las siguientes condiciones particulares:

1. La propiedad intelectual de los procesos descritos y analizados en el presente trabajo, pertenece por entero a la empresa consultora representada por el Ingeniero Director del Proyecto.
2. La empresa consultora se reserva el derecho a la utilización total o parcial de los resultados de la investigación realizada para desarrollar el siguiente proyecto, bien para su publicación o bien para su uso en trabajos o proyectos posteriores, para la misma empresa cliente o para otra.
3. Cualquier tipo de reproducción aparte de las reseñadas en las condiciones generales, bien sea para uso particular de la empresa cliente, o para cualquier otra aplicación, contará con autorización expresa y por escrito del Ingeniero Director del Proyecto, que actuará en representación de la empresa consultora.
4. En la autorización se ha de hacer constar la aplicación a que se destinan sus reproducciones así como su cantidad.
5. En todas las reproducciones se indicará su procedencia, explicitando el nombre del proyecto, nombre del Ingeniero Director y de la empresa consultora.
6. Si el proyecto pasa la etapa de desarrollo, cualquier modificación que se realice sobre él, deberá ser notificada al Ingeniero Director del Proyecto y a criterio de éste, la empresa consultora decidirá aceptar o no la modificación propuesta.
7. Si la modificación se acepta, la empresa consultora se hará responsable al mismo nivel que el proyecto inicial del que resulta el añadirla.
8. Si la modificación no es aceptada, por el contrario, la empresa consultora declinará toda responsabilidad que se derive de la aplicación o influencia de la misma.

9. Si la empresa cliente decide desarrollar industrialmente uno o varios productos en los que resulte parcial o totalmente aplicable el estudio de este proyecto, deberá comunicarlo a la empresa consultora.

10. La empresa consultora no se responsabiliza de los efectos laterales que se puedan producir en el momento en que se utilice la herramienta objeto del presente proyecto para la realización de otras aplicaciones.

11. La empresa consultora tendrá prioridad respecto a otras en la elaboración de los proyectos auxiliares que fuese necesario desarrollar para dicha aplicación industrial, siempre que no haga explícita renuncia a este hecho. En este caso, deberá autorizar expresamente los proyectos presentados por otros.

12. El Ingeniero Director del presente proyecto, será el responsable de la dirección de la aplicación industrial siempre que la empresa consultora lo estime oportuno. En caso contrario, la persona designada deberá contar con la autorización del mismo, quien delegará en él las responsabilidades que ostente.

Anexo E. Publicación

Como resultado de este PFC se ha escrito el siguiente artículo:

Diego Álvarez, Víctor López, José Luis Añamuro, Jorge López de Vergara, Óscar González de Dios y Javier Aracil, “Utilization of Temporary Reservation of Path Computed Resources for Multi-Domain PCE Protocols in WDM Networks”, en *Network of the Future 2011, 2nd IFIP International Conference, November 28-30, 2011*.

Ha sido enviado y aceptado en el siguiente congreso internacional:

Network of the Future 2011
2nd IFIP International Conference
November 28-30, 2011
Université Pierre et Marie Curie, Paris

Además, ha sido seleccionado como candidato al premio *IFIP-sponsored Best Paper Award of NoF 2011*.

Se puede obtener más información del congreso en su página web:

<http://www.network-of-the-future.org/>

A continuación se muestra la documentación acreditativa de la publicación.

Fecha: Wed, 17 Aug 2011 07:53:17 -0400 [17/08/11 13:53:17 CEST]
De: NoF11 <tpc-chairs-nof11@rp.lip6.fr> 
Para: Diego Álvarez <diego.alvarez@estudiante.uam.es>
Cc: Victor Lopez <victor.lopez@uam.es>, Jorge E. López de Vergara <jorge.lopez_vergara@uam.es>, Oscar González de Dios <ogondio@tid.es>, Javier Aracil <javier.aracil@uam.es>, tpc-chairs-nof11@rp.lip6.fr
Asunto: [NoF11] Your paper #1569476907 ('Utilization of Temporary Reservation of Path Computed Resources for Multi-Domain PCE Protocols in WDM Networks')

Dear Mr. Diego Álvarez:

We are pleased to inform you that your paper #1569476907 ('Utilization of Temporary Reservation of Path Computed Resources for Multi-Domain PCE Protocols in WDM Networks') has been accepted for NoF'11. Congratulations!

The review feedbacks are appended below or can be found at <http://edas.info/showPaper.php?m=1569476907>. Please address the issues raised by the reviewers and revise the paper accordingly for a final version.

Another email will follow to let you know the instructions for preparing the camera-ready copy and the copyright form. The guidelines will also be posted at <http://www.network-of-the-future.org>.

















Please be advised that the final acceptance of your paper for publication in the conference proceedings is contingent upon:

- 1) Your timely submission of the final manuscript (deadline Sept. 15, 2011) revised in accordance with the reviewers' comments and formatted according to the guidelines posted at <http://www-phare.lip6.fr/nof2011/>
- 2) In order to reduce the likelihood of no shows, manuscripts will be scheduled for oral presentation and published in IEEE Xplore only if at least one of the authors registers for the Conference and pays the registration fee by Sept. 15, 2011.

We received a large number of papers, and based on a rigorous review process, good quality papers have been rejected due to the lack of space. The acceptance ratio is 47%. The conference information will be updated shortly to include the full technical program and detailed instruction for preparing your presentation.

On behalf of the TPC team, we would like to use this opportunity to thank you for submitting your work to NoF'11 conference. We look forward to welcoming you to Paris in November 2011.

Sincerely,
Luis Costa and Stefano Secci, TPC co-chairs
Otto Duarte and Farouk Kamoun, conference co-chairs

Property	Change Add	Value																																										
Conference and track		2011 International Conference on the Network of the Future - Second International Conference on Network of the Future																																										
Authors		<table><tr><th>Name</th><th>ID</th><th>Flag</th><th>Affiliation</th><th>Email</th><th>Country</th></tr><tr><td>Diego Álvarez</td><td>652075</td><td></td><td>Universidad Autonoma de Madrid</td><td>diego.alvarez@estudiante.uam.es</td><td>Spain</td></tr><tr><td>Víctor Lopez</td><td>201291</td><td></td><td>Universidad Autonoma de Madrid</td><td>victor.lopez@uam.es</td><td>Spain</td></tr><tr><td>Jose Luis Añamuro</td><td>619751</td><td></td><td>Universidad Autónoma de Madrid</td><td>jose.annamuro@estudiante.uam.es</td><td>Spain</td></tr><tr><td>Jorge E. López de Vergara</td><td>130998</td><td></td><td>Universidad Autónoma de Madrid (UAM)</td><td>jorge.lopez_vergara@uam.es</td><td>Spain</td></tr><tr><td>Oscar González de Dios</td><td>193742</td><td></td><td>Telefonica I+D</td><td>ogondio@tid.es</td><td>Spain</td></tr><tr><td>Javier Aracil</td><td>3656</td><td></td><td>Universidad Autonoma de Madrid</td><td>javier.aracil@uam.es</td><td>Spain</td></tr></table>	Name	ID	Flag	Affiliation	Email	Country	Diego Álvarez	652075		Universidad Autonoma de Madrid	diego.alvarez@estudiante.uam.es	Spain	Víctor Lopez	201291		Universidad Autonoma de Madrid	victor.lopez@uam.es	Spain	Jose Luis Añamuro	619751		Universidad Autónoma de Madrid	jose.annamuro@estudiante.uam.es	Spain	Jorge E. López de Vergara	130998		Universidad Autónoma de Madrid (UAM)	jorge.lopez_vergara@uam.es	Spain	Oscar González de Dios	193742		Telefonica I+D	ogondio@tid.es	Spain	Javier Aracil	3656		Universidad Autonoma de Madrid	javier.aracil@uam.es	Spain
Name	ID	Flag	Affiliation	Email	Country																																							
Diego Álvarez	652075		Universidad Autonoma de Madrid	diego.alvarez@estudiante.uam.es	Spain																																							
Víctor Lopez	201291		Universidad Autonoma de Madrid	victor.lopez@uam.es	Spain																																							
Jose Luis Añamuro	619751		Universidad Autónoma de Madrid	jose.annamuro@estudiante.uam.es	Spain																																							
Jorge E. López de Vergara	130998		Universidad Autónoma de Madrid (UAM)	jorge.lopez_vergara@uam.es	Spain																																							
Oscar González de Dios	193742		Telefonica I+D	ogondio@tid.es	Spain																																							
Javier Aracil	3656		Universidad Autonoma de Madrid	javier.aracil@uam.es	Spain																																							
Presenter(s)		presenter not specified																																										
Title		Utilization of Temporary Reservation of Path Computed Resources for Multi-Domain PCE Protocols in WDM Networks																																										
Abstract		In recent years, Path Computation Element (PCE) architecture has been standardized as a suitable solution for path computation in multi-domain network scenarios. The Traffic Engineering Database (TED) of the PCE is updated with the information of the control plane. However, there is a delay from the PCE replies to a request, until the TED is updated with the network status information. This delay includes not only the control plane delay, but also Path Computation Element Protocol (PCEP) information exchange. This desynchronization in the TED and the real status information leads to an extra blocking situation when the Label Switch Router (LSR) tries to reserve a path, but it has been previously reserved by other LSR. To solve such problem, a recent draft is submitted to the IETF proposing new PCEP extensions for a pre-reservation of the computed path resources for a certain period. This work implements in C three multi-domain algorithms: Per-domain Path Computation, Backward-Recursive PCE-based Computation and Hierarchical Path Computation Element and it assesses their performance with and without mechanisms to reduce this extra blocking probability due to the uncertainty of the TED information.																																										
Keywords		multi-domain networks, PCE, OSPF-TE, WSON																																										
Topics		- Performance and Experiment; - GMPLS, Ethernet Carrier Grade, Burst Switching, WDM Systems; - Optical Communications and Networking																																										
Session		The program is not yet visible (tpc)																																										
DOI																																												
Status		accepted																																										
Copyright form		US September 6, 2011 11:26:29 CEST																																										
Review manuscript		<table><tr><th>Document (show)</th><th>Pages</th><th>File size</th><th>Changed</th><th>MD5</th><th>Check format / Report problem</th><th>Similarity score</th><th>Delete</th></tr><tr><td></td><td>5</td><td>437,865</td><td>July 11, 2011 18:37:20 CEST</td><td>4f02286b80c78af5dea5139f3feac84d</td><td></td><td></td><td></td></tr></table> <p>notembedded One or more fonts are not embedded. See EDAS FAQ.</p>	Document (show)	Pages	File size	Changed	MD5	Check format / Report problem	Similarity score	Delete		5	437,865	July 11, 2011 18:37:20 CEST	4f02286b80c78af5dea5139f3feac84d																													
Document (show)	Pages	File size	Changed	MD5	Check format / Report problem	Similarity score	Delete																																					
	5	437,865	July 11, 2011 18:37:20 CEST	4f02286b80c78af5dea5139f3feac84d																																								

Fecha: Tue, 30 Aug 2011 10:17:27 -0400 [30/08/11 16:17:27 CEST]
De: NoF'11 <stefano.secci@lip6.fr> 
Para: Diego Álvarez <diego.alvarez@estudiante.uam.es>
Cc: Víctor Lopez <victor.lopez@uam.es>, Jorge E. López de Vergara <jorge.lopez_vergara@uam.es>, Oscar González de Dios <ogondio@tid.es>, Javier Aracil <javier.aracil@uam.es>, tpc-chairs-nof11@rp.lip6.fr
Responder-A: tpc-chairs-nof11@rp.lip6.fr
Asunto: [NoF'11] Your paper #1569476907 (Utilization of Temporary Reservation of Path Computed Resources for Multi-Domain PCE Protocols in WDM Networks)

Dear Authors:

your paper #1569476907 (Utilization of Temporary Reservation of Path Computed Resources for Multi-Domain PCE Protocols in WDM Networks) has been short-listed for the IFIP-sponsored Best Paper Award of NoF 2011.
 Congratulations!

The final decision will be based on the paper presentation in Paris.

Regards,
 Stefano Secci
 Luis Costa
 NoF'11 TPC co-chairs

AWARDS

IFIP sponsors **one Best Paper Award** (500€).

Candidates for the Best Paper Award:

- *Utilization of Temporary Reservation of Path Computed Resources for Multi-Domain PCE Protocols in WDM Networks*
 by Diego Álvarez, Víctor Lopez, Jose Luis Añamuro, Jorge E. López de Vergara (Universidad Autónoma de Madrid, Spain); Oscar González de Dios (Telefonica I+D, Spain); Javier Aracil (Universidad Autonoma de Madrid, Spain)
- *CHANET: A Content-Centric Architecture for IEEE 802.11 MANETs*
 by Marica Amadeo, Antonella Molinaro (University Mediterranea of Reggio Calabria, Italy)
- *Secure Deployment of Application-Tailored Protocols in Future Networks*
 by Oliver Hanka (Technische Universität München, Germany); Hans Wippel (Karlsruhe Institute of Technology, Germany)
- *A Description Language for Communication Services of Future Network Architectures*
 by Rahamatullah Khondoker, Eric MSP Veith, Paul Mueller (University of Kaiserslautern, Germany)
- *A Future-Proof Application-to-Network Interface*
 by Denis Martin, Hans Wippel (Karlsruhe Institute of Technology, Germany); Helge Backhaus (Universität Karlsruhe, Germany)

The choice on the Best Paper Award will be done accordingly to the paper presentation by the TPC.

Por último, se adjunta la versión final del artículo.

Utilization of Temporary Reservation of Path Computed Resources for Multi-Domain Path Computation Element Protocols in WDM Networks

Diego Álvarez*, Víctor López*[†], José Luis Añamuro*, Jorge López de Vergara*,
Óscar González de Dios[†] and Javier Aracil*

*High Performance Computing and Networking group, Universidad Autónoma de Madrid, Spain. E-mail: victor.lopez@uam.es

[†]Telefónica I+D, Madrid, Spain, E-mail: ogondio@tid.es

Abstract—In recent years, Path Computation Element (PCE) architecture has been standardized as a suitable solution for path computation in multi-domain network scenarios. The Traffic Engineering Database (TED) of the PCE is updated with the information of the control plane. However, there is a delay from the PCE replies to a request, until the TED is updated with the network status information. This delay includes not only the control plane delay, but also Path Computation Element Protocol (PCEP) information exchange. This desynchronization in the TED and the real status information leads to an extra blocking situation when the Label Switch Router (LSR) tries to reserve a path, but it has been previously reserved by other LSR. To solve such problem, a recent draft is submitted to the IETF proposing new PCEP extensions for a pre-reservation of the computed path resources for a certain period.

This work implements in C three multi-domain algorithms: Per-domain Path Computation, Backward-Recursive PCE-based Computation and Hierarchical Path Computation Element and it assesses their performance with and without mechanisms to reduce this extra blocking probability due to the uncertainty of the TED information.

I. INTRODUCTION

With the advent of new Internet services, WDM networks are the solution to absorb such kind of high speed services that require low delay and high bandwidth utilization. The research community has done a great effort in the last years to provide a common control plane by standardizing Generalized Multi-Protocol Label Switching (GMPLS). GMPLS allows a dynamic and distributed configuration of the optical layer. However, the computation of optical paths becomes complex in terms of computation when the impairments induced by optical technologies are taken into account. If such computation is done into the GMPLS controller, the hardware requirements increase and, consequently, the node cost. In order to alleviate that issue, a Path Computation Element (PCE) architecture has been standardized [1].

In a PCE architecture, there is a PCE in each domain, which receives the request from the Path Computation Clients (PCC). As it is a single point of failure in the network, multiple PCEs can be located in the same domain [2]. PCE architecture fits with the requirements for multi-domain WDM scenarios [3]. The PCEs in a multi-domain scenario can cooperate as peers

or in a hierarchical model [2]. There are three main PCE-based algorithms for multi-domain scenarios: Per-domain Path Computation [4], Backward-Recursive PCE-based Computation (BRPC) [5] and Hierarchical Path Computation Element (H-PCE) [6]. In the recent years, the research community has worked on the PCE architecture to improve the multi-domain path computation. Authors in [7] provide an overview of the developments in the area of PCE-based traffic engineering in GMPLS networks, analyze the BRPC approach in multi-domain networks in detail, and compare its performance with non-PCE existing solutions. Authors in [8] compare the performance of BRPC and per-domain algorithms, concluding that BRPC improves per-domain algorithm in terms of blocking probability. The behavior of the H-PCE is validated in [9] and the authors assess the computation time of the H-PCE but not its blocking probability. This work implements and compares these three multi-domain protocols in terms of blocking probability.

The PCE requires the network state information, which is stored in the Traffic Engineering Database (TED). This information is updated via OSPF Link-State Advertisement (LSA) messages [2]. When the TED information is different from the network state, the PCE can reply with resources that are already reserved. When the Label Switch Router (LSR) requests the path, the control plane denies the request since these resources are occupied. Let us call this type of block as “stolen-lambda” block. A recent draft is proposed in the IETF [10] to eliminate this kind of block using PCEP extensions to pre-reserve resources. To the best of the authors knowledge, this is the first work that implements such extensions and we show the avoidance of the “stolen-lambda” block thanks to this mechanism. A PCE Proactive scheme is proposed in [11], which is similar to IETF draft [10]. The PCE Proactive scheme does not support timers for the reservation like [10] and its validation in [11] is just done by simulation.

The remaining is organized as follows: Section II provides the basics of the multi-domain PCE algorithms. Section III defines the impairment-aware routing used in this work. The evaluation between the algorithms with the implemented extensions is presented in section IV. Finally, section V con-

cludes the paper.

II. MULTI-DOMAIN PATH COMPUTATION ELEMENT PROTOCOLS

The PCE is “an entity (component, application, or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints” [1]. Path Computation Element Protocol (PCEP) follows a request/response scheme, where a PCC asks for routes to the PCE. The PCE is in charge of carrying out the route computation by taking into account the physical restrictions and the available resources in the optical layer. The PCE is aware of the network state information thanks to the flooding information mechanisms used in GMPLS. For networks without a control plane, the PCE can be located as part of the management plane [2]. When client equipment requests a new path, the request is sent via the user-to-network interface (UNI) as in traditional GMPLS networks. However, the GMPLS controller of the optical equipment redirects the request to the PCC.

One of the main motivations behind PCEs deployment is to tackle the problem of multi-domain Label Switched Paths (LSPs) establishment. There are three different computation methods namely Per-domain Path Computation, Backward-Recursive PCE-based Computation (BRPC) and Hierarchical PCE (H-PCE). Each method is discussed in the following.

1) *Per-domain Path Computation*: At this approach, the path is computed during the signalling process domain by domain. Each PCE computes the path from its ingress to egress router in its domain [4]. Consequently, the sequence of domains to be traversed must be known beforehand by the PCE in the source domain. However, there is not a mechanism to choose the best domains from the source to the destination. In addition, this procedure provides suboptimal paths because if there are multiple connections between the domains, the PCE may provide a path that is optimal locally, but not overall.

2) *Backward-Recursive PCE-based Computation*: This procedure is based on communication and cooperation between PCEs to compute optimal interdomain paths. The BRPC method starts at the destination domain, which sends to its neighbor a tree of potential paths from every ingress node to the destination node. Each PCE in the domain sequence adds its own paths from its ingress nodes to the tree and passes it to the previous domain. This process continues until the source domain is reached, which selects the best end-to-end path. Fig. 1 depicts three connected domains with one PCE per domain. Using BRPC, the PCE1 sends a request to the PCE2, which forwards it to the PCE3. The PCE3 replies with the distance from its ingress nodes (N and O) with the domain 2. The PCE2 carries out the same operation sending a tree with the possible combinations between the edge nodes from domain 1 to 3. When multiple domains are interconnected such information exchange can be complicated. If the sequence of domains is known, this process is easier. Border Gateway Protocol (BGP) can be used for this purpose. Unfortunately, BRPC does not scale with complex multi-domain topologies.

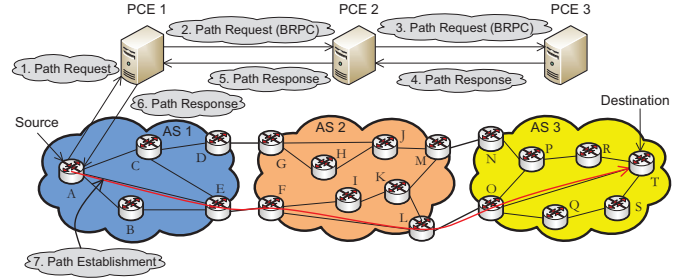


Fig. 1. Backward recursive path computation

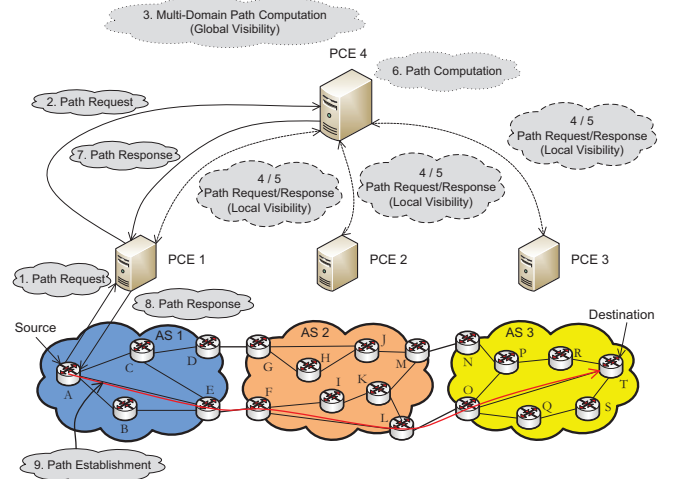


Fig. 2. H-PCE multi-domain path computation

3) *Hierarchical PCE*: Fig. 2 shows an example of a H-PCE architecture. In this architecture there is a parent PCE and some child PCEs, and they are organized in multiple levels [6]. The parent PCE does not have information of the whole network, but is only aware of the connectivity among the domains and provide coordination to the child PCEs. The path request is sent to the parent PCE, which selects a set of candidate domain paths and sends requests to the child PCEs responsible for these domains. Then the parent PCE selects the best solution and it is transmitted to the source PCE. This hierarchical model fits with the model for the Automatic Switched Optical Network (ASON), since the networks are composed by sub-networks and the routing areas have relationship between peers.

III. RWA-ALGORITHM WITH IMPAIRMENT AWARENESS

The following well-known algorithms are used in the Routing and Wavelength Assignment (RWA) units, namely Dijkstra’s k -shortest path and first-fit (FF). Both are very simple in terms of their implementation. The cost used in the routing problem may be set to distance, number of hops, link load, etc. such that depending on the scenario a route is determined based on a particular cost. The FF algorithm chooses the first available wavelength until there are no more resources available.

The characteristics of the optical network elements such as fiber and nodes should be made available to the entity that makes the IA-RWA decisions (here: the PCE). The most important linear impairments are optical signal to noise ratio (OSNR), residual chromatic dispersion (CD), polarization mode dispersion (PMD) and, in case of lightpaths, hop and technology-dependent penalties due to filter cascading at intermediate nodes [12]. The latter also depends on whether or not walk-off is present in the transmitter's wavelength. Furthermore, the optical fiber infers several non-linear impairments such as phase noise, self-phase and cross-phase modulation, and four-wave mixing. These are not always noticeable because the effects are highly depended on the used bit rate and modulation format.

To the best of the authors' knowledge, commercially available routers or sub-modules may only provide real-time data on the wavelength, and received power and OSNR of an active optical channel. Even though many research papers exist on the real-time measurement of several impairments, considering cumulated impacts is widely accepted by means of applying margins on the required OSNR and/or on the experienced non-linear phase shift. In this work, the expected OSNR and PMD values are considered to be of key importance while all other effects are covered by a single OSNR penalty in dB.

The ITU-T standard G.680 provides a detailed equation to calculate the received OSNR of a lightpath that takes into account individual properties of each span per wavelength [13]. A well-known simplified version is as follows

$$OSNR = P_{out} - \alpha L - 10 \log_{10} N + 58 - F, \quad (1)$$

with the OSNR in dB, P_{out} as the signal power in dBm at the output of the last amplifier, α as the attenuation factor in dB/km, L as the length of the fiber span in km, N as the total number of spans, and F the optical amplifier's noise factor in dB. The value of 58 in (1) equals to $10 \log_{10}(hf_i \Delta f_0)$ with h as Planck's constant, f_i as the optical frequency in Hz, and Δf_0 as the reference bandwidth in Hz. Accordingly, the PCE evaluates the following OSNR-condition for a path coming from the Routing-unit

$$OSNR > OSNR_{min} + OSNR_{imp}, \quad (2)$$

with $OSNR_{min}$ as the minimum required OSNR-value for error-free detection and $OSNR_{imp}$ as the cumulated penalty for other (non)-linear impairments. All parameters in (1) are known so the maximum number of spans N_{max} fulfilling (2) can be determined.

Regarding the experienced PMD, the equation in [13] includes the PMD values of all optical network elements constituting the lightpath. However, [14] shows a simplified version that is used to check the physical constraints namely

$$PMD = B \sqrt{\sum_{k=1}^N D_{PMD}^2 \cdot L}, \quad (3)$$

with B as the channel symbol rate, and D_{PMD} as the PMD value of each fiber span. Typically, less than 10% pulse

broadening ($a = 0.1$) is recommended for a lightpath, and therefore (3) is evaluated on being smaller or equal than a . Similar to the OSNR-limit, the value for N_{max} can also be determined for the PMD-limit and the smallest number should be stored.

It is clear that several assumptions have been made to justify the simplifications applied in (1) and (3). Regarding the optical amplifiers, the gain G fully compensates the transmission losses of a span ($G = \alpha L$), gain-control is in place such that the gain received by an individual channel agrees with the former, and all amplifiers have an equal noise factor. Regarding the fiber spans, it is assumed that these are all of equal length with average values for the attenuation factor and PMD value. Regarding the factor $10 \log_{10}(hf_i \Delta f_0)$, the difference between wavelengths on the blue or the red-side of a wavelength comb is negligible such that the value 58 in (1) is valid for all channels.

To sum up, the only lightpath-dependent parameter in (1) and (3) is N which reduces the process to check the OSNR and PMD limits to determining N for a path and to compare this value with N_{max} . Future PCE implementations may consider a more exact treatment of the physical impairments.

IV. EXPERIMENTAL RESULTS

We have run the experiments using the Chinese core network [15], which is shown in Fig. 3. For this study the maximum number of wavelengths is set to $M = 80$ and $K = 5$ in the RWA algorithm. Regarding the physical impairments, $P_{out} = 4$ dBm, $\alpha = 0.35$ dB/Km, $L = 80$ km, $OSNR_{imp} = 1.5$ dB, $OSNR_{min} = 12$ dB, $B = 10$ Gbps and $D_{PMD} = 0.2 \cdot 10^{-12}$ s/km^{0.5}. An $OSNR_{imp}$ margin is taken as the cumulated penalty of all other impairments which is similar to OSNR margins found in data sheets of commercial 10-Gigabit transceivers.

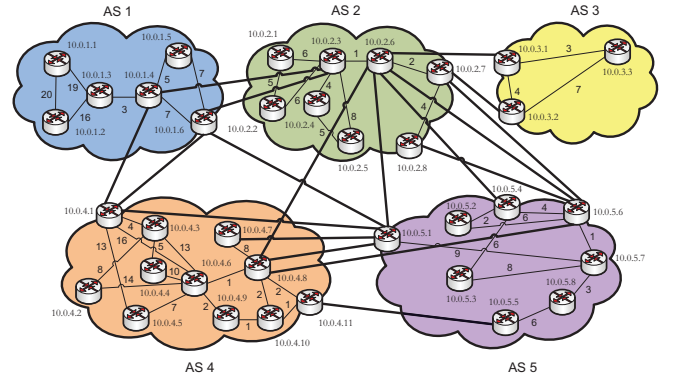


Fig. 3. Backbone Chinese topology [15]

For the experimental results, we have defined a PCC for each domain, which is requesting the routes to the PCE in its domain. For this work, we have implemented in C the PCEP protocol, the extensions to support the three multi-domain methods as well as the pre-reservation mechanism. Call requests for each node pair follow a Poisson process

for the arrival rate and the holding time. We have run 20000 requests to achieve stable results. The PCEs and PCCs are in the same server (4-core Intel(R) Xeon(R) CPU E5345 at 2.33 GHz and 8GB of RAM memory). In this light, the delays in our experiment are the PCEP protocol and the RWA algorithm with physical impairments defined in section III. There is a propagation delay (the PCEs are located in different places), which is not included in our study. The TED information is updated by the PCC when they receive the response. Depending on the PCE location, the number of control plane hops and the updating timer for the LSA messages, this time can be important and it adds an uncertainty to the path computation. This effect is out of the scope of this work.

Regarding the sequence of domains for the route computation, the BRPC and H-PCE use K-shortest path. K is higher for H-PCE case because it has a more detailed information about the connections between domains than in BRPC.

A. Impact of “stolen-lambda” block

Fig. 4 shows the blocking probability of the BRPC algorithm ($K = 4$) and the reason of the block: (1) there is no wavelength available, (2) the physical restrictions are not fulfilled and (3) the “stolen-lambda” block. The “stolen-lambda” block appears when the PCE is replying to two requests with a route which shares resources. The first request reserves the resources, but when the second request tries to reserve them they are no longer available. In this work, the TED is updated by the PCCs when they receive the response. In a real network, the propagation and the control plane delay increase the differences between the network state and the PCE TED, thus incrementing the blocking probability.

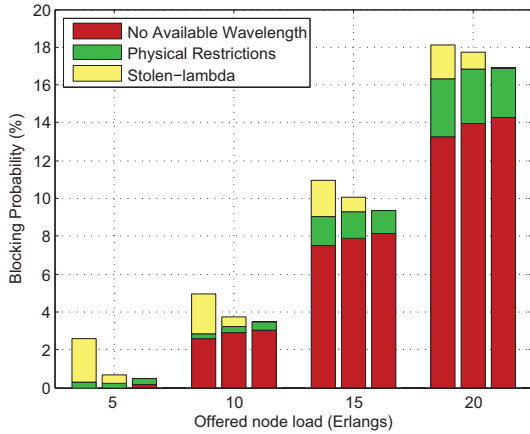


Fig. 4. Blocking probability of the BRPC with First-Fit (left), Round-Robin First-Fit (middle) and Pre-Reservation (right) algorithms

When using the RWA algorithm with FF (Fig. 4 left-bars), there is an important contribution of the “stolen-lambda” effect. To reduce this “stolen-lambda” block, we used a modification of First-Fit called Round-Robin First-Fit (RR-FF). As the PCE is using the FF mechanism, it starts the assignment process from lambda 0. Instead of starting always from lambda

0, if the PCE responses with a lambda n , RR-FF begins the search from lambda $n + 1$ in the next request. Fig. 4 (middle-bars) shows the blocking probability of the RWA with RR-FF. This mechanism is simple, it does not add complexity to the algorithm and it reduces the blocking probability due to the “stolen-lambda” effect, but not eliminates it. Pre-Reservation (PR) algorithm results are discussed in the next section (Fig. 4 right-bars).

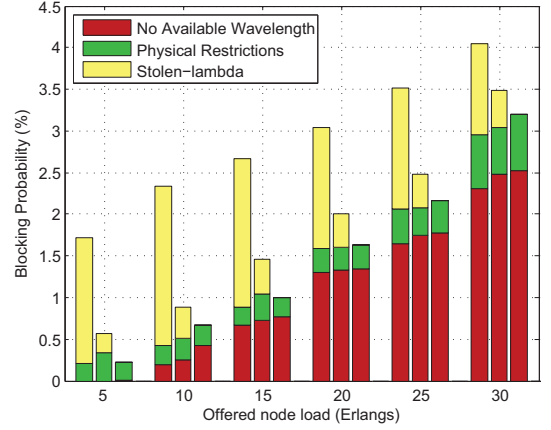


Fig. 5. Blocking probability of the H-PCE with First-Fit (left), Round-Robin First-Fit (middle) and Pre-Reservation (right) algorithms

H-PCE achieves a lower blocking probability than BRPC when using $K = 20$ (Fig. 5). The time consumption per request is higher in the H-PCE than in the BRPC. Fig. ?? shows the average delay per request for different K values using FF. The computation algorithm in the H-PCE is carried out by the parent PCE and it sends requests to the child PCEs sequentially to obtain the path segments. Then, it concatenates them and calculates the optimal path. This process implies a higher time consumption as it increases K , but the results in terms of blocking probability are better. The “stolen-lambda” block is more important in the H-PCE case than in the BRPC case (Fig. 4) in relation with the total blocking probability. For lower values of K , the time per request is very small thus reducing the “stolen-lambda” block or even eliminated at low load scenarios for BRPC and H-PCE.

B. Performance with temporary reservation of computed resources

Authors in [10] propose a temporary reservation of the resources to avoid the “stolen-lambda” block. This mechanism pre-reserves the resources in the PCE TED when there is a path request for a given time (T_{res}). Once the T_{res} timer expires, the PCE removes the reservation state of such resources. If the path was established, the OSPF LSA messages update the PCE TED properly. When using PR mechanism, the “stolen-lambda” block is eliminated if T_{res} is long enough ($T_{res} = 3s$ in our experiments). Blocking probability of PR mechanism is shown in Fig. 4 and Fig. 5 (right-bars). There is an important reduction of the blocking probability, when using the PR mechanism. The “stolen-lambda” block is completely

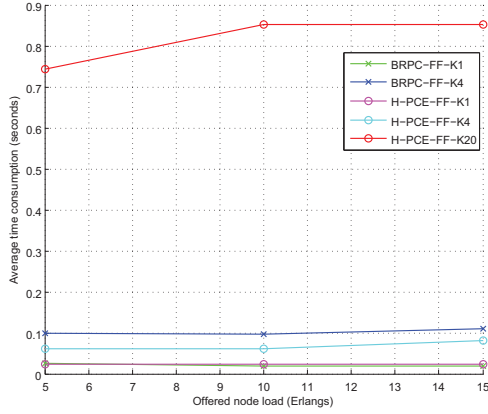


Fig. 6. BRPC and H-PCE average time consumption per request using FF

eliminated. There is a slight increment of the block because there are no more available wavelengths. The reason is that as the requests are not rejected because of the “stolen-lambda” effect, they occupy resources thus incrementing the blocking probability. Let us remark that if as there are differences from the network state information and the TED, the timer T_{res} should be incremented to obtain similar results.

Fig. 7 shows the blocking probability for per-domain, BRPC and H-PCE methods using temporary reservation. In light of this, we can conclude that H-PCE reduces the blocking probability in multi-domain scenarios and, thanks to the temporary reservation the “stolen-lambda” effect can be eliminated. For $K = 1$, BRPC and H-PCE only check a multi-domain path, thus achieving results similar to per-domain method.

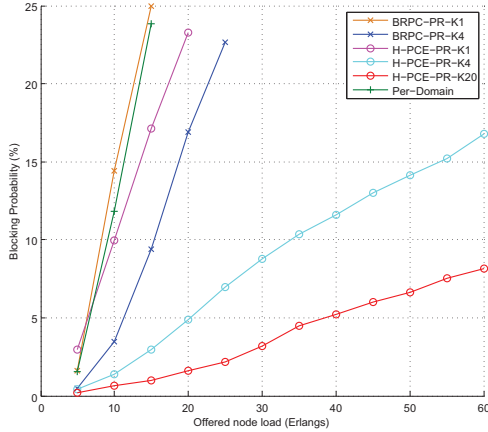


Fig. 7. Comparison in terms of blocking probability using temporal reservation

V. SUMMARY AND CONCLUSIONS

The contribution of this work is two-fold: (1) the comparison of per-domain, BRPC and H-PCE in terms of blocking probability and (2) the validation of temporary reservation mechanism as a solution to avoid “stolen-lambda” block. H-PCE has a better performance in terms of blocking probability

and its scalability is better than BRPC mechanism.

As future work, we will evaluate the algorithms behavior including the control plane uncertainty and the propagation delay between the PCEs and PCCs.

ACKNOWLEDGEMENTS

This work has been funded by the Spanish Ministry of Education and Science under project ANFORA (TEC2009-13385).

REFERENCES

- [1] A. Farrel, J.P. Vasseur, and J. Ash, “A path computation element (PCE)-based architecture,” *IETF RFC 4655*, pp. 1–40, Aug. 2006. Online (Nov. 2009): <http://tools.ietf.org/html/rfc4655>.
- [2] V. López, B. Huiszoon, J. Fernández-Palacios, O. González de Dios, and J. Aracil, “Path computation element in telecom networks: Recent developments and standardization activities,” in *Proc. Optical Networking Design and Modeling (ONDM)*, pp. 1–6, Feb. 2010, Kyoto, Japan.
- [3] M. Chamania and A. Jukan, “A survey of inter-domain peering and provisioning solutions for the next generation optical networks,” *Communications Surveys & Tutorials, IEEE*, vol. 11, no. 1, pp. 33–51, 2009.
- [4] J.P. Vasseur (Ed.), A. Ayyangar (Ed.) and R. Zhang, “A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs),” *IETF RFC 5152*, pp. 1–21, February 2008. Online: <http://tools.ietf.org/html/rfc5152>.
- [5] J.P. Vasseur (Ed.), R. Zhang, N. Bitar and J.L. Le Roux, “A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths,” *IETF RFC 5441*, pp. 1–18, April 2009. Online: <http://tools.ietf.org/html/rfc5441>.
- [6] F. Zhang, Q. Zhao, O. Gonzalez de Dios, R. Casellas and D. King, “Extensions to Path Computation Element Communication Protocol (PCEP) for Hierarchical Path Computation Elements (PCE),” *IETF Internet-Draft*, pp. 1–14, April 2011. Online: <http://tools.ietf.org/html/draft-zhang-pce-hierarchy-extensions-00>.
- [7] S. Dasgupta, J.C. de Oliveira, and J.P. Vasseur, “Path-computation-element-based architecture for interdomain MPLS/GMPLS traffic engineering: Overview and performance,” *IEEE Network*, vol. 21, no. 4, pp. 38–45, July 2007.
- [8] R. Casellas, R. Martínez, R. Muñoz, and S. Gunreben, “Enhanced Backwards Recursive Path Computation for Multi-area Wavelength Switched Optical Networks Under Wavelength Continuity Constraint,” *Journal of Optical Communications and Networking*, vol. 1, no. 2, pp. A180–A193, 2009.
- [9] R. Casellas, R. Muñoz, and R. Martínez, “Lab trial of multi-domain path computation in gmpls controlled wson using a hierarchical pce,” in *Optical Fiber Communication Conference*. Optical Society of America, 2011.
- [10] O. Gonzalez de Dios, R. Casellas and F.J. Jimenez, “PCEP Extensions for Temporary Reservation of Computed Path Resources and Support for Limited Context State in PCE,” *IETF RFC draft*, pp. 1–15, Oct 2010. Work in progress (Oct. 2010): <http://tools.ietf.org/html/draft-gonzalezdedios-pce-reservation-state-00>.
- [11] A. Giorgetti, F. Cugini, N. Sambo, F. Paolucci, N. Andrioli, and P. Castoldi, “Path state-based update of PCE traffic engineering database in wavelength switched optical networks,” *Communications Letters, IEEE*, vol. 14, no. 6, pp. 575–577, June 2010.
- [12] A.H. Gnauck, P.J. Winzer, C. Dorrer, and S. Chandrasekhar, “Linear and nonlinear performance of 42.7-Gb/s single-polarization RZ-DQPSK format,” *IEEE Photon. Technol. Lett.*, vol. 18, no. 7, pp. 883–885, April 1 2006.
- [13] ITU-T, “Physical transfer functions of optical network elements,” *G.680 Recommendation*, pp. 1–68, Jul. 2007. Online (Nov. 2009): <http://www.itu.int/rec/T-REC-G.680/en>.
- [14] J. Strand, A.L. Chiu, and R. Tkach, “Issues for routing in the optical layer,” *IEEE Commun. Mag.*, vol. 39, no. 2, pp. 81–87, Feb. 2001.
- [15] Y. Zhao, J. Zhang, Y. Ji, and W. Gu, “Routing and Wavelength Assignment Problem in PCE-Based Wavelength-Switched Optical Networks,” *Optical Communications and Networking, IEEE/OSA Journal of*, vol. 2, no. 4, pp. 196–205, 2010.